EIGHTH CANADIAN EDITION

THE ECONOMICS OF MONEY, BANKING, AND FINANCIAL MARKETS

Frederic S. Mishkin | Apostolos Serletis

Brief Contents

PART 1 Introduction 1

1 Why Study Money, Banking, and Financial Markets? 2
2 An Overview of the Financial System 22
3 What Is Money? 48

PART 2 Financial Markets 64

4 The Meaning of Interest Rates 65
5 The Behaviour of Interest Rates 87
6 The Risk and Term Structure of Interest Rates 118
7 The Stock Market, the Theory of Rational Expectations, and the Efficient Market Hypothesis 143

PART 3 Financial Institutions 164

8 An Economic Analysis of Financial Structure 165
9 Economic Analysis of Financial Regulation 189
10 Banking Industry: Structure and Competition 208
11 Financial Crises 243

PART 4 The Management of Financial Institutions 272

12 Banking and the Management of Financial Institutions 273
13 Risk Management with Derivatives 301

PART 5 Central Banking and the Conduct of Monetary Policy 329

14 Central Banks and the Bank of Canada 331
15 The Money Supply Process 355
16 Tools of Monetary Policy 381
17 The Conduct of Monetary Policy: Strategy and Tactics 424

PART 6 International Finance and Monetary Policy 453

18 The Foreign Exchange Market 454
19 The International Financial System 477

PART 7 Monetary Theory 504

20 Quantity Theory, Inflation, and the Demand for Money 505
21 The *IS* Curve 523
22 The Monetary Policy and Aggregate Demand Curves 542
23 Aggregate Demand and Supply Analysis 557
24 Monetary Policy Theory 592

25 The Role of Expectations in Monetary Policy 617
26 Transmission Mechanisms of Monetary Policy 636

Supplemental Chapters on Pearson eText

S1 Financial Crises in Emerging-Market Economies
S2 Nonbank Finance
S3 The *ISLM* Model

Contents in Detail

Pearson's Commitment to Diversity, Equity, and Inclusion xxx

Preface xxxi

PART 1 Introduction 1

CHAPTER 1
Why Study Money, Banking, and Financial Markets? 2

1.1 Why Study Financial Markets? 2
- Debt Markets and Interest Rates 3
- The Stock Market 3

1.2 Why Study Financial Institutions and Banking? 5
- Structure of the Financial System 6
- Banks and Other Financial Institutions 6
- Financial Innovation 6
- Financial Crises 7

1.3 Why Study Money and Monetary Policy? 7
- Money and Business Cycles 7
- Money and Inflation 8
- Money and Interest Rates 10
- Conduct of Monetary Policy 10
- Fiscal Policy and Monetary Policy 11

1.4 Why Study International Finance? 12
- The Foreign Exchange Market 13
- The International Financial System 14

1.5 Money, Banking, and Financial Markets and Your Career 14

1.6 How We Will Study Money, Banking, and Financial Markets 15

Concluding Remarks 16

Summary 16 • Questions 16 • Applied Problems 17 • Data Analysis Problems 18

APPENDIX TO CHAPTER 1
Defining Aggregate Output, Income, the Price Level, and the Inflation Rate 19

Aggregate Output and Income 19

Real Versus Nominal Magnitudes 19

Aggregate Price Level 20

Growth Rates and the Inflation Rate 21

CHAPTER 2
An Overview of the Financial System 22

2.1 Function of Financial Markets 22

2.2 Structure of Financial Markets 25
- Debt and Equity Markets 25
- Primary and Secondary Markets 25
- Exchanges and Over-the-Counter Markets 26
- Money and Capital Markets 27

2.3 Financial Market Instruments 27
Money Market Instruments 27

Following the Financial News Money Market Rates 28

Capital Market Instruments 29

Following the Financial News Capital Market Interest Rates 30

2.4 Internationalization of Financial Markets 31

Global Are U.S. Capital Markets Losing Their Edge? 32

International Bond Market, Eurobonds, and Eurocurrencies 32
World Stock Markets 33

2.5 Function of Financial Intermediaries: Indirect Finance 33

Following the Financial News Foreign Stock Market Indexes 34

Transaction Costs 34

Global The Importance of Financial Intermediaries Relative to Securities Markets: An International Comparison 35

Risk Sharing 36
Asymmetric Information: Adverse Selection and Moral Hazard 36
Economies of Scope and Conflicts of Interest 38

2.6 Types of Financial Intermediaries 38
Depository Institutions 38
Contractual Savings Institutions 40
Investment Intermediaries 41

2.7 Regulation of the Financial System 42
Increasing Information Available to Investors 42
Ensuring the Soundness of Financial Intermediaries 43
Financial Regulation Abroad 44

Summary 45 • Questions 46 • Applied Problems 47 • Data Analysis Problems 47

CHAPTER 3
What Is Money? 48

3.1 Meaning of Money 48

3.2 Functions of Money 49
Medium of Exchange 49
Unit of Account 50
Store of Value 51

3.3 Evolution of the Payments System 52
Commodity Money 52
Fiat Money 52
Cheques 52
Electronic Payment 53
E-Money 53

FYI Are We Headed for a Cashless Society? 54

APPLICATION Will Bitcoin or Other Cryptocurrencies Become the Money of the Future? 54

3.4 Measuring Money 55
The Bank of Canada's Monetary Aggregates 56

FYI Where Are All the Dollars? 57

Money as a Weighted Aggregate 59

Summary 61 • Questions 61 • Applied Problems 62 • Data Analysis Problems 62

PART 2 Financial Markets 64

CHAPTER 4
The Meaning of Interest Rates 65

4.1 Measuring Interest Rates 65
Present Value 66

APPLICATION Simple Present Value 67

APPLICATION How Much Is That Jackpot Worth? 68
Four Types of Credit Market Instruments 68
Yield to Maturity 69

APPLICATION Yield to Maturity on a Simple Loan 69

APPLICATION Yield to Maturity and the Yearly Payment on a Fixed-Payment Loan 71

APPLICATION Yield to Maturity and Bond Price for a Coupon Bond 72

APPLICATION Yield to Maturity on a Perpetuity 74

APPLICATION Yield to Maturity on a Discount Bond 75

4.2 The Distinction Between Interest Rates and Returns 76

Global Negative Interest Rates? Japan First, Then the United States, Then Europe 77

Maturity and the Volatility of Bond Returns: Interest-Rate Risk 79
Summary 80

4.3 The Distinction Between Real and Nominal Interest Rates 81

APPLICATION Calculating Real Interest Rates 82

Summary 84 • Questions 84 • Applied Problems 85 • Data Analysis Problems 86

CHAPTER 5
The Behaviour of Interest Rates 87

5.1 Determinants of Asset Demand 87
Wealth 88
Expected Returns 88
Risk 88
Liquidity 89
Theory of Portfolio Choice 89

5.2 Supply and Demand in the Bond Market 90
Demand Curve 90
Supply Curve 91

Market Equilibrium 92
Supply and Demand Analysis 93

5.3 Changes in Equilibrium Interest Rates 93
Shifts in the Demand for Bonds 93
Shifts in the Supply of Bonds 97

APPLICATION Changes in the Interest Rate Due to a Change in Expected Inflation: The Fisher Effect 99

APPLICATION Changes in the Interest Rate Due to a Business Cycle Expansion 101

APPLICATION Explaining Current Low Interest Rates in Europe, Japan, and the United States: Low Inflation and Secular Stagnation 102

5.4 Supply and Demand in the Market for Money: The Liquidity Preference Framework 103

5.5 Changes in Equilibrium Interest Rates in the Liquidity Preference Framework 106
Shifts in the Demand for Money 106
Shifts in the Supply of Money 106

APPLICATION Changes in the Equilibrium Interest Rate Due to Changes in Income, the Price Level, or the Money Supply 107
Changes in Income 108
Changes in the Price Level 108
Changes in the Money Supply 108

5.6 Money and Interest Rates 109

APPLICATION Does a Higher Rate of Growth of the Money Supply Lower Interest Rates? 111

Summary 114 • Questions 114 • Applied Problems 115 • Data Analysis Problems 116

CHAPTER 6
The Risk and Term Structure of Interest Rates 118

6.1 Risk Structure of Interest Rates 118
Default Risk 119

FYI Conflicts of Interest at Credit-Rating Agencies and the Global Financial Crisis 122

APPLICATION The Coronavirus Pandemic and the Baa–Treasury Spread in the United States 122
Liquidity 123
Income Tax Considerations 124
Summary 124

6.2 Term Structure of Interest Rates 125

Following the Financial News Yield Curves 126

Expectations Theory 127
Segmented Markets Theory 130
Liquidity Premium and Preferred Habitat Theories 131

FYI The Yield Curve as a Forecasting Tool for Inflation and the Business Cycle 135

Evidence on the Term Structure 135
Summary 135

APPLICATION Interpreting Yield Curves, 1986–2020 136

APPLICATION Using the Term Structure to Forecast Interest Rates 137

APPLICATION Forward Rate 139

Summary 140 • Questions 140 • Applied Problems 142
• Data Analysis Problems 142

CHAPTER 7
The Stock Market, the Theory of Rational Expectations, and the Efficient Market Hypothesis 143

7.1 Computing the Price of Common Stock 143
The One-Period Valuation Model 144
The Generalized Dividend Valuation Model 145
The Gordon Growth Model 145

7.2 How the Market Sets Stock Prices 146

APPLICATION Monetary Policy and Stock Prices 148

APPLICATION The Coronavirus Stock Market Crash of 2020 148

7.3 The Theory of Rational Expectations 148
Formal Statement of the Theory 150
Rationale Behind the Theory 150
Implications of the Theory 151

7.4 The Efficient Market Hypothesis: Rational Expectations in Financial Markets 152
Rationale Behind the Hypothesis 153
Random-Walk Behaviour of Stock Prices 154

Global Should Foreign Exchange Rates Follow a Random Walk? 155

APPLICATION Practical Guide to Investing in the Stock Market 155
How Valuable Are Reports Published by Investment Advisers? 155
Should You Be Skeptical of Hot Tips? 156

FYI Should You Hire an Ape as Your Investment Adviser? 157

Do Stock Prices Always Rise When There Is Good News? 157
Efficient Market Prescription for the Investor 157

7.5 Why the Efficient Market Hypothesis Does Not Imply That Financial Markets Are Efficient 158

APPLICATION What Do Stock Market Crashes Tell Us About the Efficient Market Hypothesis and the Efficiency of Financial Markets? 159

7.6 Behavioural Finance 159

Summary 160 • Questions 161 • Applied Problems 162
• Data Analysis Problems 163

PART 3 Financial Institutions 164

CHAPTER 8
An Economic Analysis of Financial Structure 165

8.1 Basic Facts About Financial Structure Throughout The World 165
8.2 Transaction Costs 168
How Transaction Costs Influence Financial Structure 168
How Financial Intermediaries Reduce Transaction Costs 169
8.3 Asymmetric Information: Adverse Selection and Moral Hazard 170
8.4 The Lemons Problem: How Adverse Selection Influences Financial Structure 170
Lemons in the Stock and Bond Markets 171
Tools to Help Solve Adverse Selection Problems 172

FYI The Enron Implosion 173

8.5 How Moral Hazard Affects the Choice Between Debt and Equity Contracts 176
Moral Hazard in Equity Contracts: The Principal–Agent Problem 176
Tools to Help Solve the Principal–Agent Problem 177
8.6 How Moral Hazard Influences Financial Structure in Debt Markets 179
Tools to Help Solve Moral Hazard in Debt Contracts 179
Summary 181

APPLICATION Financial Development and Economic Growth 182

FYI The Tyranny of Collateral 183

APPLICATION Is China a Counterexample to the Importance of Financial Development? 184

Summary 185 • Questions 186 • Applied Problems 187 • Data Analysis Problems 188

CHAPTER 9
Economic Analysis of Financial Regulation 189

9.1 Asymmetric Information and the Government Safety Net 189
Government Safety Net 189

Global The Spread of Government Deposit Insurance Throughout the World: Is It a Good Thing? 191

Drawbacks of the Government Safety Net 191
9.2 Types of Financial Regulation 193
Restrictions on Asset Holdings 193
Capital Requirements 194
Prompt Corrective Action 195
Financial Supervision: Chartering and Examination 195

Global Where Is the Basel Accord Heading After the Global Financial Crisis? 196

Assessment of Risk Management 197
Disclosure Requirements 198
Consumer Protection 199
Restrictions on Competition 199
Summary 199

Global International Financial Regulation 200

9.3 CDIC Deposit Insurance Coverage 203
Differential Premiums 203
Opting Out 204
Resolution Authority and the Bail-in Power 205

Summary 205 • Questions 205 • Applied Problems 206 • Data Analysis Problems 207

CHAPTER 10
Banking Industry: Structure and Competition 208

10.1 Historical Development of the Canadian Banking System 208

FYI The Dual Banking System in the United States 210

The Free Banking Experiment 210
The Provincial Notes Act, 1866 211
The Dominion Notes Act, 1870 211
The First Bank Act, 1871 212
The Bank Act, 1881–1913 212
The Finance Act, 1914 213

10.2 Financial Innovation and the Growth of the "Shadow Banking System" 213
Responses to Changes in Demand Conditions: Interest-Rate Volatility 214
Responses to Changes in Supply Conditions: Information Technology 215
Securitization and the Shadow Banking System 218
Avoidance of Existing Regulations 219

FYI Bruce Bent and the Money Market Mutual Fund Panic of 2008 221

Financial Innovation and the Decline of Traditional Banking 221

10.3 Structure of the Canadian Chartered Banking Industry 224
Schedule I, Schedule II, and Schedule III Banks 224
Competition and Technology 225

10.4 Comparison with the United States 225
Restrictions on Branching 227
Response to Branching Restrictions 227

10.5 Competition Across All Four Pillars 228
Convergence 228
Implications for Financial Consolidation 229
Separation of Banking and Other Financial Services Industries Throughout the World 229

FYI The Global Financial Crisis and the Demise of Large, Free-Standing Investment Banks 230

10.6 The Near Banks: Regulation and Structure 231
Trust Companies 231
Loan Companies 231
Cooperatives: Credit Unions and *Caisses Populaires* 232
Government Savings Institutions 233

10.7 International Banking 233

Global Ironic Birth of the Eurodollar Market 234

Eurocurrencies Market 234
Canadian Banking Overseas 235
Foreign Banks in Canada 235

10.8 The 2001 Bank Act Reform 237
Bank Holding Companies 237
Permitted Investments 238
Ownership Rules 238
The CP Act and Access to the Payments and Clearance System 239
Merger Review Policy 239
The National Financial Services OmbudService 240
Implications for the Canadian Banking Industry 240

Summary 240 • Questions 241 • Data Analysis Problems 242

CHAPTER 11
Financial Crises 243

11.1 What Is a Financial Crisis? 243

11.2 Dynamics of Financial Crises 244
Stage One: Initial Phase 244
Stage Two: Banking Crisis 246
Stage Three: Debt Deflation 247

APPLICATION The Mother of All Financial Crises: The Great Depression in the United States 248
Stock Market Crash 248
Bank Panics 248
Continuing Decline in Stock Prices 249
Debt Deflation 249
International Dimensions 250

11.3 The Global Financial Crisis of 2007–2009 250
Causes of the 2007–2009 Financial Crisis 250

FYI Collateralized Debt Obligations (CDOs) 251

Effects of the 2007–2009 Financial Crisis 252

FYI Was the Fed to Blame for the Housing Price Bubble? 253

Global The European Sovereign Debt Crisis 256

Height of the 2007–2009 Financial Crisis 257
Government Intervention and the Recovery 258

Global Worldwide Government Bailouts During the 2007–2009 Financial Crisis 258

APPLICATION Could the Coronavirus Pandemic Have Led to a Financial Crisis? ... 259

11.4 Canada and the 2007–2009 Financial Crisis 260

FYI Canada's Asset-Backed Commercial Paper Crisis 260

Subprime Mortgages in Canada 261
Why Was Canada Exempt from the Financial Crisis? 262

11.5 Response of Financial Regulation 263
Macroprudential Versus Microprudential Supervision 263
Dodd-Frank Wall Street Reform and Consumer Protection Act of 2010 264
11.6 Too Big to Fail and Future Regulation 266
What Can Be Done About the Too-Big-to-Fail Problem? 266
Beyond Dodd-Frank: Where Might Regulation Head in the Future? 267

Summary 269 • Questions 269 • Data Analysis Problems 270

PART 4 The Management of Financial Institutions 272

CHAPTER 12
Banking and the Management of Financial Institutions 273

12.1 The Bank Balance Sheet 273
Liabilities 273
Assets 275
12.2 Basic Banking 277
12.3 General Principles of Bank Management 280
Liquidity Management and the Role of Reserves 280
Asset Management 283
Liability Management 284
Capital Adequacy Management 285

APPLICATION Strategies for Managing Bank Capital 287

APPLICATION How a Capital Crunch Caused a Credit Crunch During the Global Financial Crisis 288
12.4 Managing Credit Risk 288
Screening and Monitoring 289
Long-Term Customer Relationships 290
Loan Commitments 291
Collateral and Compensating Balances 291
Credit Rationing 291
12.5 Managing Interest-Rate Risk 292
Gap and Duration Analysis 293

APPLICATION Strategies for Managing Interest-Rate Risk 294
12.6 Off-Balance-Sheet Activities 294
Loan Sales 295
Generation of Fee Income 295
Trading Activities and Risk Management Techniques 295

Global Barings, Daiwa, Sumitomo, Société Générale, and JP Morgan Chase: Rogue Traders and the Principal–Agent Problem 296

Summary 297 • Questions 298 • Applied Problems 299 • Data Analysis Problems 300

CHAPTER 13
Risk Management with Derivatives 301

13.1 Hedging 301
13.2 Interest-Rate Forward Contracts 302
APPLICATION Hedging with Interest-Rate Forward Contracts 302
Pros and Cons of Forward Contracts 303
13.3 Financial Futures Contracts and Markets 304
APPLICATION Hedging with Financial Futures 305
Organization of Trading in Financial Futures Markets 307
The Globalization of Financial Futures Markets 307
Explaining the Success of Futures Markets 308
APPLICATION Hedging Foreign Exchange Risk 309
Hedging Foreign Exchange Risk with Forward Contracts 310
Hedging Foreign Exchange Risk with Futures Contracts 310
13.4 Options 311
Stock Options 312
Futures Options 314
APPLICATION Hedging with Futures Options 317
Factors Affecting Option Premiums 318
Summary 319
13.5 Swaps 320
Interest-Rate Swap Contracts 320
APPLICATION Hedging with Interest-Rate Swaps 321
Advantages of Interest-Rate Swaps 321
Disadvantages of Interest-Rate Swaps 322
Financial Intermediaries in Interest-Rate Swaps 322
13.6 Credit Derivatives 323
Credit Options 323
Credit Swaps 323
Credit-Linked Notes 324
APPLICATION Lessons from the Global Financial Crisis: When Are Financial Derivatives Likely to Be a Worldwide Ticking Time Bomb? 324
Summary 326 • Questions 326 • Applied Problems 327 • Data Analysis Problems 328

PART 5 Central Banking and the Conduct of Monetary Policy 329

CHAPTER 14
Central Banks and the Bank of Canada 331

14.1 Origins of the Bank of Canada 331
Global Establishment of Selected Central Banks 332
14.2 Formal Structure of the Bank of Canada 332
The Functions of the Bank of Canada 333

FYI The Political Environment and the Bank of Canada 333

FYI Role of the Bank's Research Staff 336

14.3 How Independent Is the Bank of Canada? 336
The Changing Face of the Bank of Canada 338
14.4 Should the Bank of Canada Be Independent? 341
The Case for Independence 341

FYI Economics and Politics 342

The Case Against Independence 342
Central Bank Independence and Macroeconomic Performance Throughout the World 343
14.5 Explaining Central Bank Behaviour 343
14.6 Structure and Independence of Foreign Central Banks 344
Federal Reserve System 344

FYI The Special Role of the Federal Reserve Bank of New York 347

FYI U.S. Presidential Attacks on the Independence of the Fed 348

European Central Bank 349
Bank of England 351
Bank of Japan 351
The Trend Toward Greater Independence 352

Summary 352 • Questions 353 • Data Analysis Problems 354

CHAPTER 15
The Money Supply Process 355

15.1 Three Players in the Money Supply Process 355
15.2 The Bank of Canada's Balance Sheet 356
Liabilities 356
Assets 357
15.3 Control of the Monetary Base 358
Bank of Canada Open Market Operations 358
Shifts from Deposits into Currency 359
Loans to Financial Institutions 360
Other Factors That Affect the Monetary Base 361
Overview of the Bank of Canada's Ability to Control the Monetary Base 362
15.4 Multiple Deposit Creation: A Simple Model 363
Deposit Creation: The Single Bank 363
Deposit Creation: The Banking System 364
Deriving the Formula for Multiple Deposit Creation 367
Critique of the Simple Model 368

15.5 Factors That Determine the Money Supply 369
Changes in the Nonborrowed Monetary Base, MB_n 369
Changes in Borrowed Reserves, BR, from the Bank of Canada 369
Changes in the Desired Reserve Ratio, r_d 369
Changes in Excess Reserves 370
Changes in Currency Holdings 370
15.6 Overview of the Money Supply Process 370
15.7 The Money Multiplier 371
Deriving the Money Multiplier 371
Intuition Behind the Money Multiplier 373
Money Supply Response to Changes in the Factors 374
APPLICATION Quantitative Easing and the Money Supply During the Global Financial and the Coronavirus Crises 375
Summary 377 • Questions 378 • Applied Problems 379 • Data Analysis Problems 380

CHAPTER 16
Tools of Monetary Policy 381
16.1 The Framework for the Implementation of Monetary Policy 381
The Large Value Transfer System (LVTS) 382
Non-LVTS (ACSS) Transactions 383
The Bank of Canada's Policy Rate 383
The Operating Band for the Overnight Interest Rate 384
FYI Monetary Policy Implementation in the LVTS Environment 385
The Bank of Canada's Standing Facilities 386
The Bank of Canada's Implementation of the Operating Band for the Overnight Interest Rate 387
16.2 The Market for Settlement Balances and the Channel/Corridor System for Setting the Overnight Interest Rate 387
Demand and Supply in the Market for Reserves 387
Demand Curve 387
Supply Curve 388
Market Equilibrium 389
APPLICATION How the Bank of Canada's Operating Procedures Limit Fluctuations in the Overnight Interest Rate 389
16.3 The Bank of Canada's Approach to Monetary Policy 390
FYI Price-Level Targeting Versus Inflation-Rate Targeting 391
How Monetary Policy Affects the Economy 391
Nominal Interest Rates and Monetary Policy 393
16.4 Conventional Monetary Policy Tools 394
Open Market (Buyback) Operations 395
Settlement Balances Management 397
APPLICATION Monetary Control in the Channel/Corridor System 400
Bank of Canada Lending 401
Operation of the Standing Lending Facility 401

FYI Emergency Lending Assistance to Troubled Banks 404

16.5 Nonconventional Monetary Policy Tools and Quantitative Easing in the Wake of the Global Financial Crisis and the Coronavirus Pandemic 405
Liquidity Provision 405

FYI Bank of Canada Programs During the Covid-19 Pandemic 406

Large-Scale Asset Purchases 407
Quantitative Easing Versus Credit Easing 407

FYI Quantitative Easing in the Channel/Corridor System: The Floor System of Monetary Control 408

Forward Guidance 411
Negative Interest Rates on Banks' Deposits 412

FYI Helicopter Money 412

16.6 Monetary Policy Tools of the Federal Reserve and the European Central Bank 413
Policy Tools of the Federal Reserve 413
Monetary Policy Tools of the European Central Bank 418

Summary 420 • Questions 421 • Applied Problems 422 • Data Analysis Problems 423

CHAPTER 17
The Conduct of Monetary Policy: Strategy and Tactics 424

17.1 The Price Stability Goal and the Nominal Anchor 424
The Role of a Nominal Anchor 425
The Time-Inconsistency Problem 425

17.2 Other Goals of Monetary Policy 426
High Employment and Output Stability 426
Economic Growth 427
Stability of Financial Markets 427
Interest-Rate Stability 427
Stability in Foreign Exchange Markets 428

17.3 Should Price Stability Be the Primary Goal of Monetary Policy? 428
Hierarchical Versus Dual Mandates 428
Price Stability as the Primary, Long-Run Goal of Monetary Policy 429

17.4 Inflation Targeting 429
Inflation Targeting in New Zealand, Canada, and the United Kingdom 430

Global The European Central Bank's Monetary Policy Strategy 432

Advantages of Inflation Targeting 433
Disadvantages of Inflation Targeting 434

17.5 Lessons for Monetary Policy Strategy from the Global Financial Crisis 436
Implications for Inflation Targeting 437

FYI The Fed's New Monetary Policy Strategy: Average Inflation Targeting 438

17.6 Should Central Banks Try to Stop Asset-Price Bubbles? 439
Two Types of Asset-Price Bubbles 439
The Debate over Whether Central Banks Should Try to Pop Bubbles 440

17.7 Tactics: Choosing the Policy Instrument 443
Criteria for Choosing the Policy Instrument 445

17.8 Tactics: The Taylor Rule 446

FYI The Bank of Canada's Use of the Taylor Rule 448

FYI Bank of Canada Watchers 449

Summary 449 • Questions 450 • Applied Problems 451
• Data Analysis Problems 451

PART 6 International Finance and Monetary Policy 453

CHAPTER 18
The Foreign Exchange Market 454

18.1 Foreign Exchange Market 454

Following the Financial News Foreign Exchange Rates 455

What Are Foreign Exchange Rates? 455
Why Are Exchange Rates Important? 455
How Is Foreign Exchange Traded? 456

18.2 Exchange Rates in the Long Run 457
Theory of Purchasing Power Parity 457

APPLICATION Burgernomics: Big Macs and PPP 459
Factors That Affect Exchange Rates in the Long Run 461

18.3 Exchange Rates in the Short Run: A Supply and Demand Analysis 463
Supply Curve for Domestic Assets 463
Demand Curve for Domestic Assets 464
Equilibrium in the Foreign Exchange Market 465

18.4 Explaining Changes in Exchange Rates 465
Shifts in the Demand for Domestic Assets 465
Recap: Factors That Change the Exchange Rate 468

APPLICATION Effects of Changes in Interest Rates on the Equilibrium Exchange Rate 470

APPLICATION The Global Financial Crisis and the U.S. Dollar 472

APPLICATION Brexit and the British Pound 473

Summary 474 • Questions 475 • Applied Problems 476
• Data Analysis Problems 476

CHAPTER 19
The International Financial System 477

19.1 Intervention in the Foreign Exchange Market 477
Foreign Exchange Intervention and the Money Supply 477
Unsterilized Intervention 479
Sterilized Intervention 480

19.2 Balance of Payments 481
Current Account 481
Financial Account 481

Global Should We Worry About the Large U.S. Current Account Deficit? 482

19.3 Exchange Rate Regimes in the International Financial System 482
Gold Standard 483
The Bretton Woods System 483
How a Fixed Exchange Rate Regime Works 484
Speculative Attacks 486

APPLICATION The Foreign-Exchange Crisis of September 1992 486
The Policy Trilemma 488

APPLICATION How Did China Accumulate US$4 Trillion of International Reserves? 489
Monetary Unions 489
Managed Float 490

Global Will the Euro Survive? 490

19.4 Capital Controls 491
Controls on Capital Outflows 491
Controls on Capital Inflows 491

19.5 The Role of the IMF 492
Should the IMF Act as an International Lender of Last Resort? 492

19.6 International Considerations and Monetary Policy 493
Direct Effects of the Foreign Exchange Market on Monetary Policy 493
Exchange Rate Considerations 494

19.7 To Peg or Not to Peg: Exchange-Rate Targeting as an Alternative Monetary Policy Strategy 494
Advantages of Exchange-Rate Targeting 494
Disadvantages of Exchange-Rate Targeting 495
When Is Exchange-Rate Targeting Desirable for Industrialized Countries? 497
When Is Exchange-Rate Targeting Desirable for Emerging Market Countries? 498
Currency Boards 498

Global Argentina's Currency Board 499

Dollarization 499

Summary 500 • Questions 501 • Applied Problems 502 • Data Analysis Problems 502

PART 7 Monetary Theory 504

CHAPTER 20
Quantity Theory, Inflation, and the Demand for Money 505

20.1 Quantity Theory of Money 505
Velocity of Money and Equation of Exchange 505
From the Equation of Exchange to the Quantity Theory of Money 507
Quantity Theory and the Price Level 507
Quantity Theory and Inflation 508

APPLICATION Testing the Quantity Theory of Money 509

20.2 Budget Deficits and Inflation 511
Government Budget Constraint 511

FYI Modern Monetary Theory 513

Hyperinflation 513

APPLICATION The Zimbabwean Hyperinflation 514

20.3 Keynesian Theories of Money Demand 514
Transactions Motive 515
Precautionary Motive 515
Speculative Motive 515
Putting the Three Motives Together 515

20.4 Portfolio Theories of Money Demand 516
Theory of Portfolio Choice and Keynesian Liquidity Preference 516
Other Factors That Affect the Demand for Money 517
Summary 517

20.5 Empirical Evidence on the Demand for Money 518
Interest Rates and Money Demand 518
Stability of Money Demand 519

Summary 519 • Questions 520 • Applied Problems 521 • Data Analysis Problems 522

CHAPTER 21
The *IS* Curve 523

21.1 Planned Expenditure and Aggregate Demand 523

21.2 The Components of Aggregate Demand 524
Consumption Expenditure 524

FYI Meaning of the Word *Investment* 525

Planned Investment Spending 525
Government Purchases and Taxes 527
Net Exports 528

21.3 Goods Market Equilibrium 529
Solving for Goods Market Equilibrium 529
Deriving the *IS* Curve 530

21.4 Understanding the *IS* Curve 530
What the *IS* Curve Tells Us: Intuition 530
What the *IS* Curve Tells Us: Numerical Example 530
Why the Economy Heads Toward Equilibrium 532
21.5 Factors That Shift the *IS* Curve 532
Changes in Government Purchases 532
APPLICATION The Vietnam War Buildup, 1964–1969 533
Changes in Taxes 534
APPLICATION The U.S. Fiscal Stimulus Package of 2009 535
Changes in Autonomous Spending 536
Changes in Financial Frictions 538
Summary of Factors That Shift the *IS* Curve 538
Summary 538 • Questions 539 • Applied Problems 540 • Data Analysis Problems 541

CHAPTER 22
The Monetary Policy and Aggregate Demand Curves **542**
22.1 The Bank of Canada and Monetary Policy 542
22.2 The Monetary Policy Curve 543
Why the Monetary Policy Curve Has an Upward Slope 543
Shifts in the *MP* Curve 544
Movements Along Versus Shifts in the *MP* Curve 545
APPLICATION Movements Along the *MP* Curve: The Rise in the Federal Funds Rate Target, 2004–2006 and 2015–2019 546
APPLICATION Shift in the *MP* Curve: Autonomous Monetary Easing in Canada During the Global Financial and Coronavirus Crises 547
22.3 The Aggregate Demand Curve 548
Deriving the Aggregate Demand Curve Graphically 548
Factors That Shift the Aggregate Demand Curve 550
FYI Deriving the Aggregate Demand Curve Algebraically 550
Summary 554 • Questions 554 • Applied Problems 555 • Data Analysis Problems 556

CHAPTER 23
Aggregate Demand and Supply Analysis **557**
23.1 Business Cycles and Inflation 557
Business Cycles 557
Inflation 560
23.2 Aggregate Demand 561
Components of Aggregate Demand 561
Following the Financial News Aggregate Output, Unemployment, and Inflation 561

Deriving the Aggregate Demand Curve 562
Factors That Shift the Aggregate Demand Curve 562

FYI What Does *Autonomous* Mean? 563

23.3 Aggregate Supply 566
Long-Run Aggregate Supply Curve 566
Short-Run Aggregate Supply Curve 567
Price Stickiness and the Short-Run Aggregate Supply Curve 569

23.4 Shifts in the Aggregate Supply Curves 569
Shifts in the Long-Run Aggregate Supply Curve 569
Shifts in the Short-Run Aggregate Supply Curve 570

23.5 Equilibrium in Aggregate Demand and Supply Analysis 573
Short-Run Equilibrium 574
Aggregate Demand and Supply Analysis Using an Aggregate Output Index 574
How the Short-Run Equilibrium Moves to the Long-Run Equilibrium over Time 575
Self-Correcting Mechanism 578

23.6 Changes in Equilibrium: Aggregate Demand Shocks 578

APPLICATION The Bank of Canada Disinflation, 1981–1985 580

23.7 Changes in Equilibrium: Aggregate Supply (Inflation) Shocks 581

APPLICATION Negative Supply Shocks in the United States, 1973–1975 and 1978–1980 583

23.8 Conclusions from Aggregate Demand and Supply Analysis 584

APPLICATION *AD/AS* Analysis of the Great Recession of 2007–2009 in the United States 585

APPLICATION An *AD/AS* Analysis of the Covid-19 Recession in Canada 587

Summary 589 • Questions 589 • Applied Problems 590 • Data Analysis Problems 591

CHAPTER 24
Monetary Policy Theory 592

24.1 Response of Monetary Policy to Shocks 592
Response to an Aggregate Demand Shock 593
Response to a Supply Shock 594
The Bottom Line: The Relationship Between Stabilizing Inflation and Stabilizing Economic Activity 597

24.2 How Actively Should Policymakers Try to Stabilize Economic Activity? 598
Lags and Policy Implementation 599

FYI The Activist/Nonactivist Debate over the Obama Fiscal Stimulus Package in the United States 600

24.3 Inflation: Always and Everywhere a Monetary Phenomenon 600

24.4 Causes of Inflationary Monetary Policy 601
High Employment Targets and Inflation 602

24.5 Monetary Policy at the Effective Lower Bound 605
Deriving the Aggregate Demand Curve with the Effective Lower Bound 605
The Disappearance of the Self-Correcting Mechanism at the Effective Lower Bound 607

APPLICATION Nonconventional Monetary Policy and Quantitative Easing 609
Liquidity Provision 609
Asset Purchases and Quantitative Easing 609
Management of Expectations 611

APPLICATION Abenomics and the Shift in Japanese Monetary Policy in 2013 612

Summary 614 • Questions 615 • Applied Problems 616
• Data Analysis Problems 616

CHAPTER 25
The Role of Expectations in Monetary Policy 617

25.1 Lucas Critique of Policy Evaluation 617
Econometric Policy Evaluation 618

APPLICATION The Term Structure of Interest Rates 618

25.2 Policy Conduct: Rules or Discretion? 619
Discretion and the Time-Inconsistency Problem 619
Types of Rules 620
The Case for Rules 620

FYI The Political Business Cycle and Richard Nixon 621

The Case for Discretion 621
Constrained Discretion 622

Global The Demise of Monetary Targeting in Switzerland 622

25.3 The Role of Credibility and a Nominal Anchor 623
Benefits of a Credible Nominal Anchor 623
Credibility and Aggregate Demand Shocks 624
Credibility and Aggregate Supply Shocks 626

APPLICATION A Tale of Three Oil Price Shocks 627
Credibility and Anti-Inflation Policy 629

Global Ending the Bolivian Hyperinflation: A Successful Anti-Inflation Program 630

FYI Credibility and the Bank of Canada's Victory over Inflation 631

25.4 Approaches to Establishing Central Bank Credibility 631
Nominal GDP Targeting 632
Appoint "Conservative" Central Bankers 632

Summary 633 • Questions 633 • Applied Problems 634
• Data Analysis Problems 635

CHAPTER 26
Transmission Mechanisms of Monetary Policy 636

26.1 Transmission Mechanisms of Monetary Policy 637
Traditional Interest-Rate Channels 637
Other Asset Price Channels 638

Credit View 641
Why Are Credit Channels Likely to Be Important? 644

FYI Consumers' Balance Sheets and the Great Depression 644

APPLICATION The Great Recession 645
26.2 Lessons for Monetary Policy 645

APPLICATION Applying the Monetary Policy Lessons to Japan's Two Lost Decades 647

Summary 648 • Questions 648 • Applied Problems 650 • Data Analysis Problems 650

Glossary **G-1**

Guide to Commonly Used Symbols **GCUS-1**

Credits **C-1**

Index **I-1**

Supplemental Chapters on Pearson eText

SUPPLEMENTAL CHAPTER 1
Financial Crises in Emerging-Market Economies S1-1

S1.1 Dynamics of Financial Crises in Emerging-Market Economies S1-1
Stage One: Initial Phase S1-1
Stage Two: Currency Crisis S1-5
Stage Three: Full-Fledged Financial Crisis S1-6

APPLICATION Crisis in South Korea, 1997–1998 ..... S1-7
Financial Liberalization and Globalization Mismanaged S1-8
Perversion of the Financial Liberalization and Globalization Process: Chaebols and the South Korean Crisis S1-9
Stock Market Decline and Failure of Firms Increase Uncertainty S1-10
Adverse Selection and Moral Hazard Problems Worsen, and the Economy Contracts S1-11
Currency Crisis Ensues S1-11
Final Stage: Currency Crisis Triggers Full-Fledged Financial Crisis S1-11
Recovery Commences S1-13

APPLICATION The Argentine Financial Crisis, 2001–2002 ..... S1-13
Severe Fiscal Imbalances S1-13
Adverse Selection and Moral Hazard Problems Worsen S1-14
Bank Panic Begins S1-14
Currency Crisis Ensues S1-14
Currency Crisis Triggers Full-Fledged Financial Crisis S1-15
Recovery Begins S1-17

Global When an Advanced Economy Is Like an Emerging-Market Economy: The Icelandic Financial Crisis of 2008 S1-18

S1.2 Preventing Emerging-Market Financial Crises S1-18
Beef Up Prudential Regulation and Supervision of Banks S1-18
Encourage Disclosure and Market-Based Discipline S1-19

Limit Currency Mismatch S1-19
Sequence Financial Liberalization S1-20

Summary S1-20 • Questions S1-21

SUPPLEMENTAL CHAPTER 2
Nonbank Finance S2-1

S2.1 Insurance S2-1
Life Insurance S2-1
Property and Casualty Insurance S2-2
The Competitive Threat from the Banking Industry S2-4
Credit Insurance S2-4

FYI The AIG Blowup S2-4

The New Legislative Framework S2-5

FYI The Global Financial Crisis and the Monoline Insurers S2-5

APPLICATION Insurance Management S2-6
Screening S2-6
Risk-Based Premiums S2-7
Restrictive Provisions S2-7
Prevention of Fraud S2-7
Cancellation of Insurance S2-7
Deductibles S2-8
Coinsurance S2-8
Limits on the Amount of Insurance S2-8
Summary S2-8
S2.2 Pension Funds S2-9
Public Pension Plans S2-10
Private Pension Plans S2-10

FYI Should Public Pension Plans Be Privatized? S2-11

Personal Pension Plans S2-12
S2.3 Finance Companies S2-12
S2.4 Securities Market Operations S2-13
Investment Banking S2-13
Securities Brokers and Dealers S2-14
Organized Exchanges S2-15
S2.5 Mutual Funds S2-15

FYI Sovereign Wealth Funds: Are They a Danger? S2-16

Money Market Mutual Funds S2-17
S2.6 Hedge Funds S2-17
S2.7 Private Equity and Venture Capital Funds S2-18
S2.8 Government Financial Intermediation S2-19
Crown Finance Companies S2-19
Government-Sponsored Enterprises in the United States S2-20

FYI The Global Financial Crisis and the Bailout of Fannie Mae and Freddie Mac S2-20

Summary S2-21 • Questions S2-22 • Applied Problems S2-23

SUPPLEMENTAL CHAPTER 3
The *ISLM* Model S3-1

S3.1 Keynes's Fixed Price Level Assumption and the *IS* Curve S3-1
The *LM* Curve S3-1
Equilibrium in the Market for Money: The *LM* Curve S3-2
S3.2 *ISLM* Approach to Aggregate Output and Interest Rates S3-4
S3.3 Factors that Cause the *LM* Curve to Shift S3-5
S3.4 Changes in Equilibrium Level of the Interest Rate and Aggregate Output S3-7
Response to a Change in Monetary Policy S3-7
Response to a Change in Fiscal Policy S3-8
APPLICATION The Covid-19 Crisis in Canada S3-9
S3.5 Effectiveness of Monetary Versus Fiscal Policy S3-11
Monetary Policy Versus Fiscal Policy: The Case of Complete Crowding Out S3-11
APPLICATION Targeting Money Supply Versus Interest Rates S3-13
S3.6 *ISLM* Model in the Long Run S3-16
Summary S3-18 • Questions S3-19

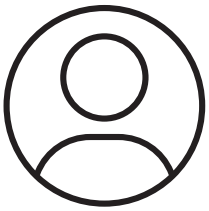
APPENDIX TO SUPPLEMENTAL CHAPTER 3
Algebra of the *ISLM* Model S3-20

Basic Closed-Economy *ISLM* Model S3-20
IS and *LM* Curves S3-21
Solution of the Model S3-21
Implications S3-21
Open-Economy *ISLM* Model S3-22
Implications S3-23

ONLINE APPENDICES IN MYLAB ECONOMICS

Chapter 4: 4.1 Measuring Interest-Rate Risk: Duration
Chapter 5: 5.1 Loanable Funds Framework
Chapter 9: 9.1 The 1980s Canadian Banking Crisis
9.2 Banking Crises Throughout the World
Chapter 12: 12.1 Duration Gap Analysis
12.2 Measuring Bank Performance
Chapter 15: 15.1 The Bank of Canada's Balance Sheet and the Monetary Base
15.2 The M2+ Money Multiplier
15.3 The Great Depression Bank Panics, 1930–1933, and the Money Supply
Chapter 17: 17.1 Monetary Targeting
17.2 A Brief History of Bank of Canada Policymaking
Chapter 23: 23.1 The Effects of Macroeconomic Shocks on Asset Prices
23.2 Aggregate Demand and Supply: A Numerical Example
Chapter 26: 26.1 Evaluating Empirical Evidence: The Debate Over the Importance of Money in Economic Fluctuations

Pearson's Commitment to Diversity, Equity, and Inclusion



Pearson is dedicated to creating bias-free content that reflects the diversity, depth, and breadth of all learners' lived experiences.

We embrace the many dimensions of diversity, including but not limited to race, ethnicity, gender, sex, sexual orientation, socioeconomic status, ability, age, and religious or political beliefs.

Education is a powerful force for equity and change in our world. It has the potential to deliver opportunities that improve lives and enable economic mobility. As we work with authors to create content for every product and service, we acknowledge our responsibility to demonstrate inclusivity and incorporate diverse scholarship so that everyone can achieve their potential through learning. As the world's leading learning company, we have a duty to help drive change and live up to our purpose to help more people create a better life for themselves and to create a better world.

Our ambition is to purposefully contribute to a world where

- everyone has an equitable and lifelong opportunity to succeed through learning
- our educational content accurately reflects the histories and lived experiences of the learners we serve
- our educational products and services are inclusive and represent the rich diversity of learners
- our educational content prompts deeper discussions with students and motivates them to expand their own learning (and worldview)

Accessibility

We are also committed to providing products that are fully accessible to all learners. As per Pearson's guidelines for accessible educational Web media, we test and retest the capabilities of our products against the highest standards for every release, following the WCAG guidelines in developing new products for copyright year 2022 and beyond.

You can learn more about Pearson's commitment to accessibility at **https://www.pearson.com/us/accessibility.html**.

Contact Us

While we work hard to present unbiased, fully accessible content, we want to hear from you about any concerns or needs with this Pearson product so that we can investigate and address them:

Please contact us with concerns about any potential bias at **https://www.pearson.com/report-bias.html**.

For accessibility-related issues, such as using assistive technology with Pearson products, alternative text requests, or accessibility documentation, email the Pearson Disability Support team at **disability.support@pearson.com**.

Pearson

Preface

There has never been a more exciting time to teach money and banking. The recent worldwide financial crisis and the coronavirus pandemic cast a spotlight on the importance of banks, financial markets, and monetary policy to the health of our economy. Rick Mishkin experienced this firsthand when he served as a governor of the Federal Reserve System from 2006 to 2008, and in this book, we emphasize the rich tapestry of recent economic events to enliven the study of money, banking, and financial markets.

WHAT'S NEW?

This text has undergone a major revision with new material in every part of the book, including updating of all data through 2020 whenever possible. However, it continues to retain the basic hallmarks that have made it the best-selling textbook on money and banking over the past seven editions in Canada. As with past editions, this eighth edition uses basic economic principles to explain financial markets, financial institutions, and monetary policy with rigour and clarity. With each edition, we update content and features based on market feedback from economics professors and students using the book as well as the latest world financial episodes. For the past several editions, the digital assets for this book, which are available on MyLab Economics, have evolved and expanded.

Compelling New Material on the Coronavirus Pandemic

The coronavirus pandemic that spread throughout the world in 2020 and 2021 is one of the signature events of the twenty-first century. This has required the addition of many timely new sections, applications, and boxes throughout the book:

- A new application that uses the analysis of the risk structure of interest rates to explain the effect of the coronavirus pandemic on the Baa-Treasury spread in the United States (Chapter 6).
- A new application on the coronavirus stock market crash of 2020 (Chapter 7), which illustrates how stock market prices are set.
- A new application on whether the coronavirus pandemic could have led to a financial crisis (Chapter 11) shows how to apply the analysis of the dynamics of financial crises to explain when financial crises might occur in the future.
- A new application on the effects of quantitative easing on the money supply during the coronavirus crisis (Chapter 15), which shows how to apply the model of the money supply process to recent data.
- An update to the section on nonconventional monetary policy tools and quantitative easing (Chapter 16) to discuss how they were used during the coronavirus pandemic.
- An updated FYI box on Bank of Canada lending facilities during the coronavirus crisis (Chapter 16).
- An update on the application discussing shifts in the *MP* curve (Chapter 22) to explain why the actions taken at the onset of the coronavirus pandemic were an autonomous monetary easing.
- A new application that shows how *AD/AS* analysis can explain what happened during the coronavirus recession (Chapter 23).

A More Real-World Approach to Monetary Theory

Part 7 of the text, Monetary Theory, has been substantially revised to bring the analysis closer to the real world by using actual data when conducting *AD/AS* analysis of business cycle episodes. Chapter 23 has been revised in two major ways. It now starts with an entirely new section, *Business Cycles*, that discusses what aggregate demand and supply analysis is trying to explain, that is, cyclical fluctuations in output, unemployment, and inflation. The *AD/AS* analysis is then developed using an aggregate output index (where 100 is potential GDP), which replaces aggregate output on the horizontal axis of the *AD/AS* diagram. This approach has two important advantages over the analysis in previous editions. First, it enables *AD/AS* diagrams to be a little simpler because the long-run aggregate supply curve does not have to be shown in the diagram because its position is always the same at an aggregate output index of 100. Second, and far more important, doing the analysis with an aggregate output index enables the *AD/AS* diagram to use actual data when it is used to describe what happened during particular business cycle episodes, such as the Great Recession and the Covid-19 recession. This change makes *AD/AS* analysis far more relevant to students because they now see that it can explain actual data and is not just a theoretical construct. This new approach is then used throughout the rest of the chapters in the monetary policy part of the book.

Additional New Material on Financial Markets and Money

Other new developments in the money and banking field have prompted us to add the following new boxes and applications that keep the text current:

- A revision of the application on whether Bitcoin or other cryptocurrencies will become the money of the future (Chapter 3) enables students to better understand the attributes of money.
- A new FYI box, which discusses U.S. presidential attacks on the independence of the Fed (Chapter 14), illustrates how central bank independence can be threatened.
- A new FYI box on the Fed's new monetary policy strategy, announced in August 2020, which involves a modification of inflation targeting (Chapter 17).
- A new FYI box on modern monetary theory (Chapter 20), which discusses the theory that argues that the Green New Deal can be easily paid for by having the central bank buy government bonds to fund the resulting large budget deficits.
- An addition of another rationale for explaining why the monetary policy curve slopes upward (Chapter 22).

FEATURES

MyLab Economics

MyLab is the teaching and learning platform that empowers you to reach every student. By combining trusted author content with digital tools and a flexible platform, MyLab personalizes the learning experience and improves results for each student. Learn more about MyLab at www.pearson.com/mylab.

Important Digital Assets

Mini-Cases These cases on MyLab Economics present unique economics scenarios and questions that provide real-life context for key chapter concepts. The solutions to the mini-cases are included in the Instructor's Solutions Manual.

Real-Time Data Figures Students can see the latest data in some in-text figures as interactives in the enhanced Pearson eText (using the Federal Reserve Bank of St. Louis's FRED database) and learn where they can access this data when they need to throughout their career.

Real-Time Data Analysis Problems Some of the text's Data Analysis Problems are available in MyLab Economics. These problems in MyLab Economics ask students to apply up-to-the-minute data, taken from the St. Louis Federal Reserve Bank's FRED database, so that they can understand what is happening in the economy in real time. These problems, marked with **Real-Time Data Analysis**, ask the student to download data and then use the data to answer questions about current issues in money and banking. In MyLab Economics, these easy-to-assign and automatically graded Real-Time Data Analysis exercises communicate directly with the FRED site so that students see updated data every time new data is posted by FRED. Thus the Real-Time Data Analysis exercises offer a no-fuss solution for instructors who want to make the most current data a central part of their macroeconomics course. These exercises will give students practice manipulating data, a skill that employers value highly.

Online Appendices and Supplemental Chapters The eighth Canadian edition of *The Economics of Money, Banking, and Financial Markets* includes 14 appendices available on MyLab Economics that cover additional topics and more technical material that instructors might want to include in their courses. Instructors can either use the appendices in class to supplement the material in the textbook or recommend them to students who want to expand their knowledge of the money and banking field. Additional supplemental chapters on financial crises in emerging-market economies, nonbank finance, and the *ISLM* model are also available within the eText on MyLab Economics. Please find them at www.pearson.com/mylab.

New Current Events Boxes Global, Following the Financial News, and FYI boxes bring currency into your classroom with author-written articles that connect key concepts with real-life current events. Our authors regularly add new or revised content to ensure that your students have relevant examples to help them engage with the course.

New Interactive Figures Interactive Figures bring concepts to life, helping students see the concepts through directed explorations and purposeful manipulation. These figures are assignable and encourage active learning, critical thinking, and conceptual understanding. These interactive graphs and figures can be found by learners within the eText and can be assigned by instructors through our MyLab Multimedia Library.

Ready-to-Go Ready-to-Go is a roadmap to planning your course. Our Ready-to-Go helps instructors unlock content and tools available within our MyLabs. Curated for each title, this website provides a path for instructors to explore and implement resources for teaching and learning. Using the framework of before, during, and after class, this online instructor's manual helps pinpoint the tools you need from the instructor support materials and online platforms to support your course planning for in-class, online, and hybrid instruction.

Dynamic Study Modules Using a highly personalized, algorithmically driven process, Dynamic Study Modules continuously assess student performance and provide additional practice in the areas where they struggle the most. Each Dynamic Study Module, accessed by computer, smartphone, or tablet, promotes fast learning and long-term retention.

SUPPLEMENTS

The following supplements are available for download from the MyLab Economics Instructor Resources page. Contact your Pearson rep for access information and instructions if you don't have a MyLab account:

- Test Banks (in Word, TestGen, and LMS formats)
- Instructor's Solutions Manual
- PowerPoint Presentations

ACKNOWLEDGMENTS

This book is the result of efforts by many people. We are extremely grateful to Kimberley Veevers, executive portfolio manager; Söğüt Y. Güleç, content manager; Rachel Stuckey, content developer; Ainsley Somerville, senior project manager; and many others at Pearson Canada who have contributed to the completion of this edition.

We are also grateful to all of the many people who commented on various editions of the book, made valuable suggestions, and kindly provided us with data. Although we have done our best to make this edition as complete and error free as possible, as most of you know, perfection is impossible. We would greatly appreciate any suggestions for improvement.

Please send your comments to serletis@ucalgary.ca.

Frederic S. Mishkin
Apostolos Serletis
2021

ABOUT THE AUTHORS

Frederic S. Mishkin is the Alfred Lerner Professor of Banking and Financial Institutions at the Graduate School of Business, Columbia University. He is also a research associate at the National Bureau of Economic Research, co-director of the U.S. Monetary Policy Forum, a member of the Squam Lake Working Group on Financial Reform, and past president of the Eastern Economics Association. Since receiving his Ph.D. from the Massachusetts Institute of Technology in 1976, he has taught at the University of Chicago, Northwestern University, Princeton University, and Columbia. He has also received an honorary professorship from the People's (Renmin) University of China. From 1994 to 1997, he was Executive Vice President and Director of Research at the Federal Reserve Bank of New York and an associate economist of the Federal Open Market Committee of the Federal Reserve System. From September 2006 to August 2008, he was a member (governor) of the Board of Governors of the Federal Reserve System.

Professor Mishkin's research focuses on monetary policy and its impact on financial markets and the aggregate economy. He is the author of more than 20 books, including *Macroeconomics: Policy and Practice,* Second Edition (Pearson, 2015); *Financial Markets and Institutions*, Ninth Edition (Pearson, 2018); *Monetary Policy Strategy* (MIT Press, 2007); *The Next Great Globalization: How Disadvantaged Nations Can Harness Their Financial Systems to Get Rich* (Princeton University Press, 2006); *Inflation Targeting: Lessons from the International Experience* (Princeton University Press, 1999); *Money, Interest Rates, and Inflation* (Edward Elgar, 1993); and *A Rational Expectations Approach to Macroeconometrics: Testing Policy Ineffectiveness and Efficient Markets Models* (University of Chicago Press, 1983). In addition, he has published more than 200 articles in such journals as *American Economic Review, Journal of Political Economy*, *Econometrica*, *Quarterly Journal of Economics, Journal of Finance*, and *Journal of Monetary Economics*.

Professor Mishkin has served on the editorial board of *American Economic Review* and has been an associate editor at *Journal of Business and Economic Statistics, Journal of Applied Econometrics, Journal of Economic Perspectives, Journal of International Money and Finance,* and *Journal of Money, Credit and Banking*; he also served as the editor of the Federal Reserve Bank of New York's *Economic Policy Review*. He is currently an associate editor (member of the editorial board) at six academic journals, including *International Finance*; *Finance India*; *Review of Development Finance; Borsa Economic Review; PSU Research Review* and *Emerging Markets;* and *Finance and Trade*. He has been a consultant to the Board of Governors of the Federal Reserve System, the World Bank, and the International Monetary Fund, as well as to many central banks throughout the world. He was also a member of the International Advisory Board to the Financial Supervisory Service of South Korea and an adviser to the Institute for Monetary and Economic Research at the Bank of Korea. Professor Mishkin was a senior fellow at the Federal Deposit Insurance Corporation's Center for Banking Research and was an academic consultant to and serves on the Economic Advisory Panel and Monetary Advisory Panel of the Federal Reserve Bank of New York.

Apostolos Serletis is professor of economics at the University of Calgary. Since receiving his Ph.D. from McMaster University in 1984, he has held visiting appointments at the University of Texas at Austin, the Athens University of Economics and Business, and the Research Department of the Federal Reserve Bank of St. Louis.

Professor Serletis's teaching and research interest focus on monetary and financial economics, macroeconometrics, and nonlinear and complex dynamics. He is the author of 12 books, including *Principles of Economics* (Canadian edition), with R. Glenn Hubbard, Anthony Patrick O'Brien, and Jason Childs; *Macroeconomics: A Modern Approach*: *First Canadian Edition*, with Robert J. Barro (Nelson, 2010); *The Demand for Money: Theoretical and Empirical Approaches* (Springer, 2007); *Financial Markets and Institutions: Canadian Edition*, with Frederic S. Mishkin and Stanley G. Eakins (Addison-Wesley, 2004); and *The Theory of Monetary Aggregation*, co-edited with William A. Barnett (Elsevier, 2000).

In addition, he has published more than 250 articles in such journals as the *Journal of Economic Literature*, *Journal of Monetary Economics*, *Journal of Money, Credit and Banking*, *Journal of Economic Dynamics and Control*, *Journal of Econometrics*, *Journal of the American Statistical Association*, *Journal of Applied Econometrics*, *Journal of Business and Economic Statistics*, *Macroeconomic Dynamics*, *Journal of Banking and Finance*, *Oxford Bulletin of Economics and Statistics*, *Canadian Journal of Economics*, *Econometric Reviews*, and *Studies in Nonlinear Dynamics and Econometrics*.

SUPPLEMENTAL CHAPTER 1

Financial Crises in Emerging-Market Economies

Learning Objectives

S1.1 Identify the key features of the three stages of a financial crisis in an emerging-market economy.

S1.2 Discuss the different measures used to prevent emerging-market financial crises.

Preview

Before 2007, the most prominent examples of severe financial crises in recent times came from abroad. Particularly vulnerable were *emerging-market economies*, which opened their markets to the outside world in the 1990s with high hopes of rapid economic growth and reduced poverty. Instead, however, many of these nations experienced financial crises as debilitating as the Great Depression was in the United States.

Most dramatic were the Mexican crisis that began in 1994, the East Asian crisis that began in July 1997, and the Argentine crisis, which started in 2001. These events presented a puzzle for economists: How can a developing country shift so dramatically from a path of high growth—as did Mexico and particularly the East Asian countries of Thailand, Malaysia, Indonesia, the Philippines, and South Korea—to such a sharp decline in economic activity?

In this chapter we apply the asymmetric information theory of financial crises developed in Chapter 11 to investigate the cause of frequent and devastating financial crises in emerging-market economies. First we explore the dynamics of financial crises in emerging-market economies. Then we apply our analysis to the events surrounding the financial crises that took place in two of these economies in recent years and explore why these events caused such devastating contractions of economic activity.

S1.1 DYNAMICS OF FINANCIAL CRISES IN EMERGING-MARKET ECONOMIES

LO S1.1 Identify the key features of the three stages of a financial crisis in an emerging-market economy.

The dynamics of financial crises in **emerging-market economies**—economies in an early stage of market development that have recently opened up to the flow of goods, services, and capital from the rest of the world—resemble those found in advanced countries but with some important differences. Figure S1-1 outlines the stages of events in financial crises in these emerging-market economies. We will address these stages in this section.

Stage One: Initial Phase

Crises in advanced economies can be triggered by a number of different factors. But in emerging-market economies, financial crises develop along two basic paths—either through the mismanagement of financial liberalization and globalization or through

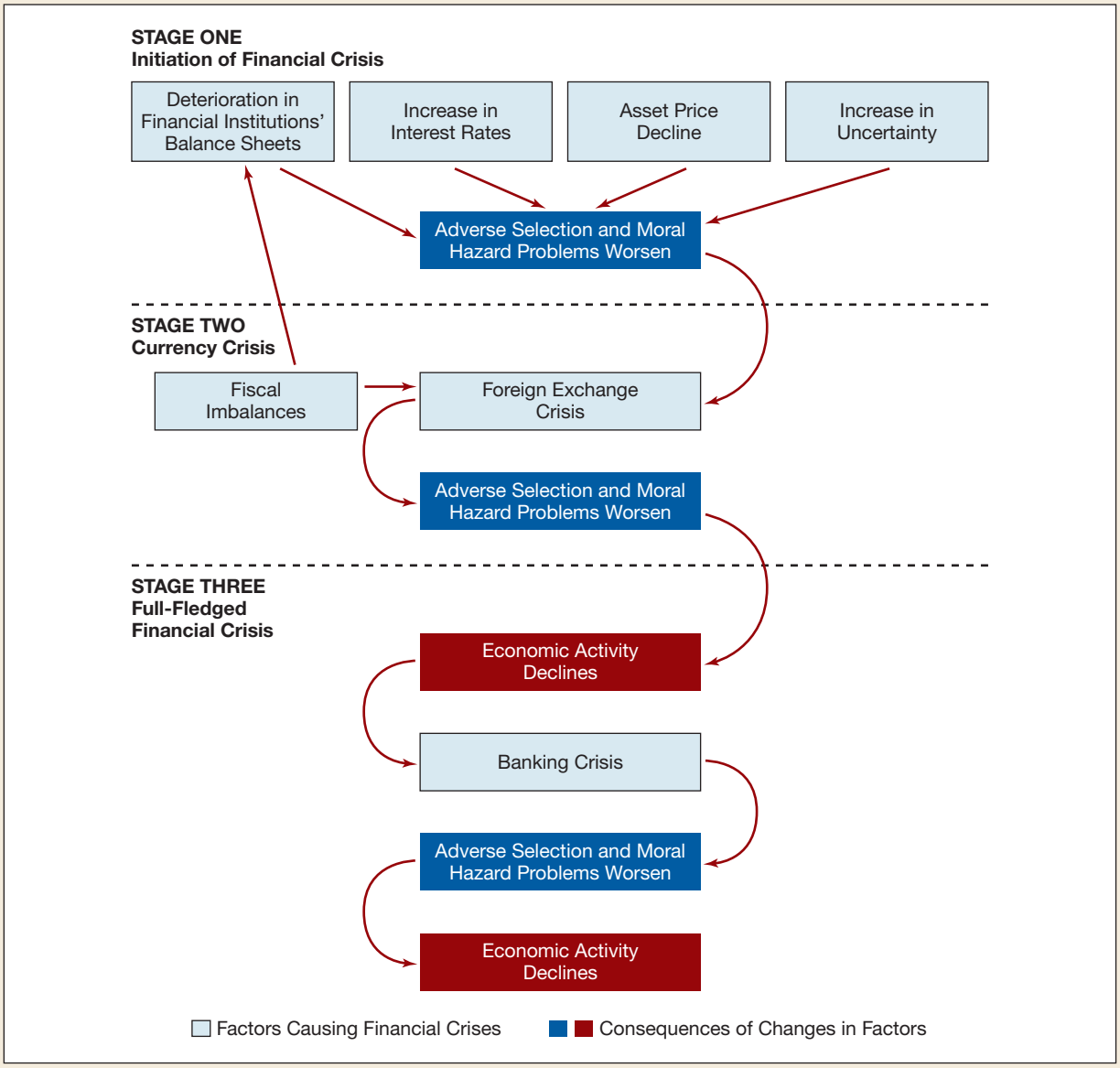


FIGURE S1-1 Sequence of Events in Emerging-Market Financial Crises
The solid arrows trace the sequence of events during a financial crisis. The sections separated by the dashed horizontal lines show the different stages of a financial crisis.

severe fiscal imbalances. The first path, mismanagement of financial liberalization and globalization, is the most common path and the one that precipitated the crises in Mexico in 1994 and in many East Asian countries in 1997.

Path A: Credit Boom and Bust The seeds of a financial crisis in an emerging-market economy are often sown when the country liberalizes its domestic financial systems by eliminating restrictions on financial institutions and markets, a process known as

financial liberalization, and opens up its economy to flows of capital and financial firms from other nations, a process called **financial globalization**. Countries often begin the process with solid fiscal policy. In the period prior to its crisis, Mexico ran a budget deficit of only 0.7% of GDP, a number to which most advanced countries would aspire. And the countries of East Asia even ran budget surpluses before their crises struck.

It is often said that emerging-market financial systems have a weak "credit culture," characterized by ineffective screening and monitoring of borrowers and lax government supervision of banks. Credit booms that accompany financial liberalization in emerging-market nations are typically marked by especially risky lending practices, sowing the seeds for enormous loan losses down the road. The financial globalization process adds fuel to the fire because it allows domestic banks to borrow abroad. Banks pay high interest rates to attract foreign capital and so are able to rapidly increase their lending. The capital inflow is further stimulated by government policies that fix the value of the domestic currency to the U.S. dollar, which provides foreign investors a sense of comfort.

Just as in advanced countries like Canada and the United States, lending booms in emerging-market economies end in lending crashes. Significant loan losses emerge from long periods of risky lending, weakening bank balance sheets and prompting banks to cut back on lending. The deterioration in bank balance sheets has an even greater negative impact on lending and economic activity than in advanced countries, which tend to have sophisticated securities markets and large nonbank financial sectors that can pick up the slack when banks falter. So as banks stop lending, there are really no other players to solve adverse selection and moral hazard problems (as shown by the arrow pointing from the first factor in the top row of Figure S1-1).

The story so far suggests that a lending boom and crash are inevitable outcomes of financial liberalization and globalization in emerging-market countries, but this is not the case. These events occur only when there is an institutional weakness that prevents the nation from successfully navigating the liberalization/globalization process. More specifically, if prudential regulation and supervision to limit excessive risk-taking are strong, the lending boom and bust will not happen. Why are regulation and supervision typically weak? The answer is the principal–agent problem, discussed in Chapter 8, in which powerful domestic business interests are encouraged by the prospect of high payoffs to pervert the financial liberalization process. Politicians and prudential supervisors are ultimately agents for voters-taxpayers (principals): That is, the goal of politicians and prudential supervisors is, or should be, to protect the taxpayers' interest. Taxpayers almost always bear the cost of bailing out the banking sector if losses occur.

Once financial markets have been liberalized, however, powerful business interests that own banks may benefit if prudential supervisors do not do their jobs properly, and so they may encourage these supervisors to act in ways that are not in the public interest. Powerful business interests that contribute heavily to politicians' campaigns are often able to persuade politicians to weaken regulations that restrict the business interests' banks from engaging in high-risk/high-payoff strategies. After all, if bank owners achieve growth and expand bank lending rapidly, they stand to make a fortune. But if a bank gets into trouble, the government is likely to bail it out, and the taxpayer will be left holding the bill. In addition, these business interests can make sure that the supervisory agencies, even in the presence of tough regulations, lack the resources they need to effectively monitor banking institutions or to close them down.

Powerful business interests have acted to prevent supervisors from doing their jobs properly in advanced countries like Canada and the United States. However, the weak

institutional environments of emerging-market countries add to the perversion of the financial liberalization process. In emerging market economies, business interests are far more powerful than they are in advanced economies, where a better-educated public and a free press monitor (and punish) politicians and bureaucrats who are not acting in the public interest. Not surprisingly, then, the cost to society of the principal–agent problem is particularly high in emerging-market economies.

Path B: Severe Fiscal Imbalances The inappropriate financing of government spending also can place emerging-market economies on a path toward financial crisis. The financial crisis in Argentina in 2001–2002 was caused by improper financing of this type; other crises—for example, the crises in Russia in 1998, in Ecuador in 1999, and in Turkey in 2001—also show some elements of this path.

When Willie Sutton, a famous bank robber, was asked why he robbed banks, he answered, "Because that's where the money is." Governments in emerging-market countries sometimes have the same attitude. When they face large fiscal imbalances and cannot finance their debt, they often cajole or force domestic banks into purchasing government debt. Investors who lose confidence in the ability of the government to repay this debt then unload the bonds, which causes their prices to plummet. Banks that hold this debt then face a big hole on the asset side of their balance sheets, which leads to a huge decline in their net worth. With less capital, these institutions have fewer resources to lend, and so lending declines. The situation is even worse if the decline in bank capital leads to a bank panic in which many banks fail at the same time. Severe fiscal imbalances therefore lead to a weakening of the banking system, which leads to a worsening of adverse selection and moral hazard problems (as shown by the arrow from the first factor in the third row of Figure S1-1).

Additional Factors Other factors often play a role in the first stage of a financial crisis. For example, a precipitating factor in some crises (such as the Mexican crisis) was a rise in interest rates caused by events abroad, such as a tightening of U.S. monetary policy. When interest rates rise, high-risk firms are the borrowers most willing to pay the high interest rates, so the adverse selection problem becomes more severe. In addition, high interest rates reduce firms' cash flows, forcing them to seek funds in external capital markets in which asymmetric information problems are greater. Thus increases in interest rates abroad, which raise domestic interest rates, can increase adverse selection and moral hazard problems (as shown by the arrow from the second factor in the top row of Figure S1-1).

Because asset markets are not as large in emerging-market countries as they are in advanced countries, they play a less prominent role in financial crises. Asset price declines in the stock market do, nevertheless, decrease the net worth of firms and so increase adverse selection problems. There is less collateral for lenders to seize, which leads to increased moral hazard problems because, given their lower net worth, the owners of the firm have less to lose if they engage in riskier activities than they engaged in before the crisis. Asset price declines therefore can worsen adverse selection and moral hazard problems directly and also indirectly, by causing deteriorations in banks' balance sheets from asset write-downs (as shown by the arrow pointing from the third factor in the first row of Figure S1-1).

As happens in advanced countries, when an emerging-market economy is in a recession or a prominent firm fails, people become more uncertain about the returns on their investment projects. In emerging-market countries, notoriously unstable political systems are another source of uncertainty. When uncertainty increases, it becomes hard for lenders

to screen good credit risks from bad ones and to monitor the activities of the firms to which they have loaned money, again worsening adverse selection and moral hazard problems (as shown by the arrow pointing from the last factor in the first row of Figure S1-1).

Stage Two: Currency Crisis

As the effects of the factors shown at the top of the diagram in Figure S1-1 build on each other, participants in the foreign exchange market sense an opportunity: They can make huge profits if they bet on a depreciation of the currency of an emerging-market country. As discussed in Chapter 19, a currency that is fixed against the U.S. dollar now becomes subject to a **speculative attack**, in which speculators engage in massive sales of the currency. As the currency sales flood the market, supply far outstrips demand, the value of the currency collapses, and a currency crisis ensues (see the Stage Two section of Figure S1-1). High interest rates abroad, increases in uncertainty, and falling asset prices all play a role. The deterioration of bank balance sheets and severe fiscal imbalances, however, are the two key factors that trigger the speculative attacks and plunge the economy into a full-scale, vicious downward spiral of currency crisis, financial crisis, and meltdown.

Deterioration of Bank Balance Sheets Triggers Currency Crises

When banks and other financial institutions are in trouble, governments have a limited number of options. Defending their currencies by raising interest rates should encourage capital inflows, but if the government raises interest rates, banks must pay more to obtain funds. This increase in costs decreases bank profitability, which may lead to insolvency. Thus, when the banking system is in trouble, the government and the central bank are now stuck between a rock and a hard place: If they raise interest rates too much, they will destroy their already weakened banks and further weaken their economy. If they don't, they won't be able to maintain the value of their currency.

Speculators in the market for foreign currency will recognize the troubles in a country's financial sector and will know when the foreign government's ability to raise interest rates and defend the currency is likely to become so costly that the government will give up and allow its currency to depreciate. They will seize an almost sure-thing bet because they know that the currency can only go downward in value. Speculators will then engage in a feeding frenzy and will sell the currency in anticipation of its decline, which will provide them with huge profits. These sales will rapidly use up the country's holdings of reserves of foreign currency because the country will have to sell its reserves to buy the domestic currency and keep it from falling in value. Once the country's central bank has exhausted its holdings of foreign-currency reserves, the cycle will end. The government no longer has the resources to intervene in the foreign exchange market and must let the value of the domestic currency fall: That is, the government must allow a devaluation.

Severe Fiscal Imbalances Trigger Currency Crises

We have seen that severe fiscal imbalances can lead to a deterioration of bank balance sheets and so can indirectly help produce a currency crisis. Fiscal imbalances can also directly trigger a currency crisis. When government budget deficits spin out of control, foreign and domestic investors begin to suspect that the country may not be able to pay back its government debt and so will start pulling money out of the country and selling the domestic currency. Recognition that the fiscal situation is out of control thus results in a speculative attack against the currency, which eventually results in its collapse.

Stage Three: Full-Fledged Financial Crisis

In contrast to most advanced economies that typically denominate debt in domestic currency, emerging-market economies denominate many debt contracts in foreign currency (usually U.S. dollars), leading to what is referred to as **currency mismatch**. An unanticipated depreciation or devaluation of the domestic currency (for example, pesos) in an emerging-market country increases the debt burden of domestic firms in terms of the domestic currency. That is, it takes more pesos to pay back the dollarized debt. Since most firms price the goods and services they produce in the domestic currency, the firms' assets do not rise in value in terms of pesos, but their debt does. The depreciation of the domestic currency increases the value of debt relative to assets, and the firms' net worth declines. The decline in net worth then increases the adverse selection and moral hazard problems described earlier. A decline in investment and economic activity follows (as shown by the Stage Three section of Figure S1-1).

We now see how the institutional structure of debt markets in emerging-market countries interacts with currency devaluations to propel the economies into full-fledged financial crises. A currency crisis, with its resulting depreciation of the domestic currency, leads to a deterioration of firms' balance sheets that sharply increases adverse selection and moral hazard problems. Economists often refer to concurrent currency and financial crises as the "twin crises."

The collapse of a currency also can lead to higher inflation. The central banks in most emerging-market countries, in contrast to those in advanced countries, have little credibility as inflation fighters. Thus, a sharp depreciation of the currency after a currency crisis leads to immediate upward pressure on import prices. A dramatic rise in both actual and expected inflation is likely to follow, which will cause domestic interest rates to increase. The resulting increase in interest payments causes reductions in firms' cash flows, which leads to increased asymmetric information problems, since firms are now more dependent on external funds to finance their investments. The increase in adverse selection and moral hazard problems then leads to a reduction in investment and economic activity.

As shown in Figure S1-1, further deterioration in the economy occurs. The collapse in economic activity and the deterioration of cash flow and firm and household balance sheets means that many debtors are no longer able to pay off their debts, resulting in substantial losses for banks. Sharp rises in interest rates also have a negative effect on banks' profitability and balance sheets. Even more problematic for the banks is the sharp increase in the value of their foreign-currency-denominated liabilities after the devaluation. Thus, bank balance sheets are squeezed from both sides—the value of their assets falls as the value of their liabilities rises.

Under these circumstances, the banking system often suffers a banking crisis in which many banks fail (as happened in the United States during the Great Depression). The banking crisis and the contributing factors in the credit markets explain a further worsening of adverse selection and moral hazard problems and a further collapse of lending and economic activity in the aftermath of the crisis.

We now apply our analysis to the study of two of the many financial crises that have struck emerging-market economies in recent years.[1] First, we examine the South

[1]For more detail on the South Korean and Argentine crises as well as a discussion of other emerging-market financial crises, see Frederic S. Mishkin, *The Next Great Globalization: How Disadvantaged Nations Can Harness Their Financial Systems to Get Rich* (Princeton, NJ: Princeton University Press, 2006).

Korean crisis of 1997–1998 because it illustrates the first path toward a financial crisis, which operates through mismanagement of financial liberalization and globalization. Second, we look at the Argentine crisis of 2001–2002, which was triggered by severe fiscal imbalances, the second path that we studied.

APPLICATION Crisis in South Korea, 1997–1998

Prior to its financial crisis in 1997, South Korea was one of the great economic success stories. In 1960, seven years after the Korean War ended, the country was still extremely poor, with an annual income per person of less than US$2000 (in today's dollars), putting South Korea on par with Somalia. During the postwar period, South Korea pursued an export-oriented strategy that helped it become one of the world's major economies. With an annual growth rate of nearly 8% from 1960 to 1997, it was one of the leaders in the "Asian miracle," a term used to refer to formerly poor Asian countries that had subsequently experienced rapid economic growth. By 1997, South Korea's income per person had risen by more than a factor of 10.

South Korea's macroeconomic fundamentals were strong before the crisis. Figure S1-2 shows that inflation was below 5% in 1996, Figure S1-3 shows that real output growth was close to 7%, and Figure S1-4 shows that unemployment was low. The government budget was enjoying a slight surplus, something that most advanced countries have been unable to achieve.

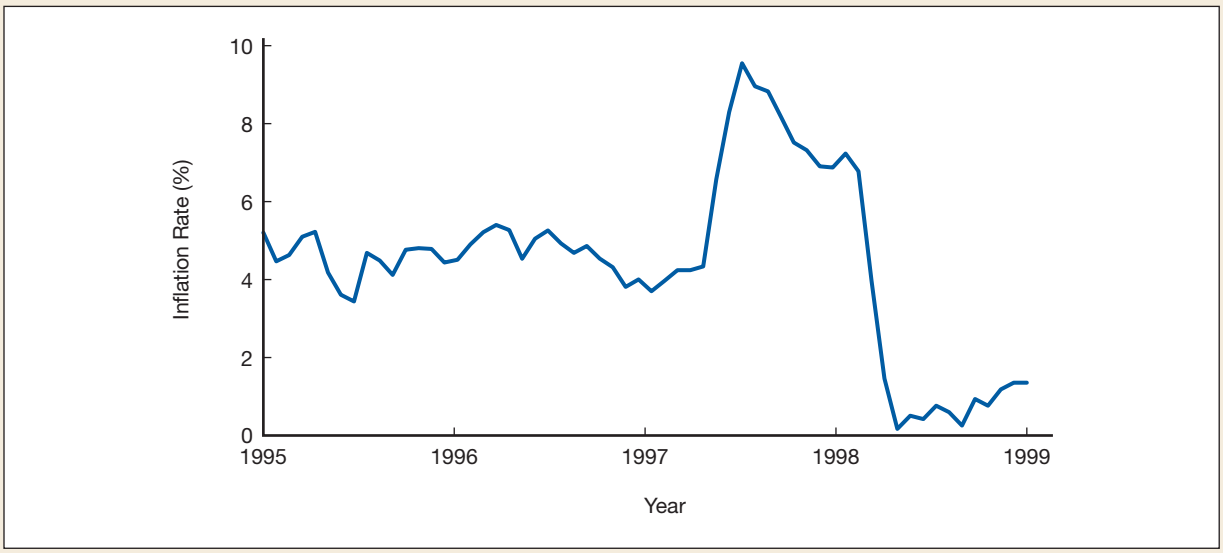


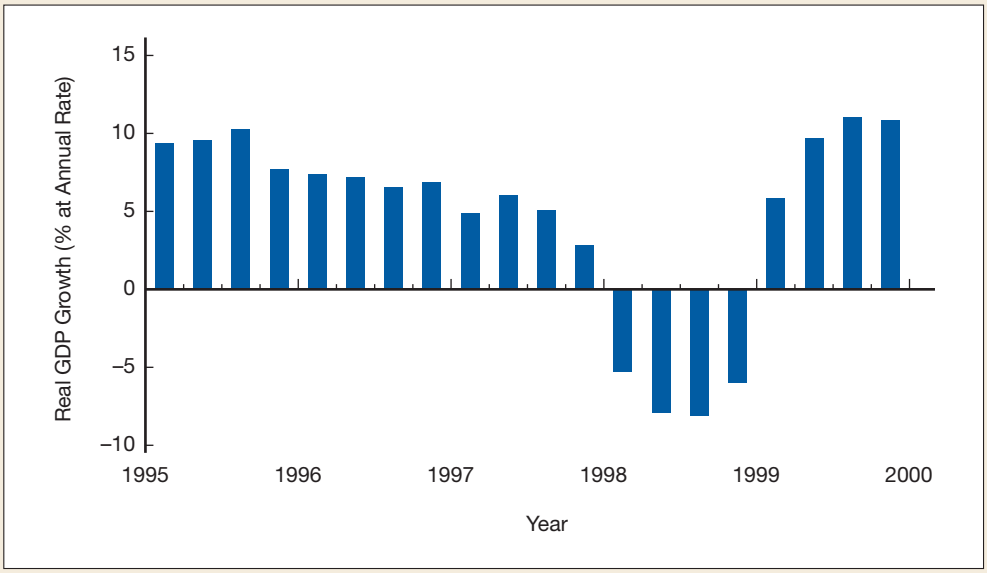
FIGURE S1-2 Inflation, South Korea, 1995–1999
Inflation was below 5% before the crisis but jumped to nearly 10% during the crisis.
Source: International Monetary Fund. *International Financial Statistics.* www.imfstatistics.org/imf/.

FIGURE S1-3

Real GDP Growth, South Korea, 1995–1999

Real GDP growth was close to 7% at an annual rate before the crisis but declined at more than a 6% rate during the crisis.

Source: International Monetary Fund. *International Financial Statistics.* www.imfstatistics.org/imf/.



Financial Liberalization and Globalization Mismanaged

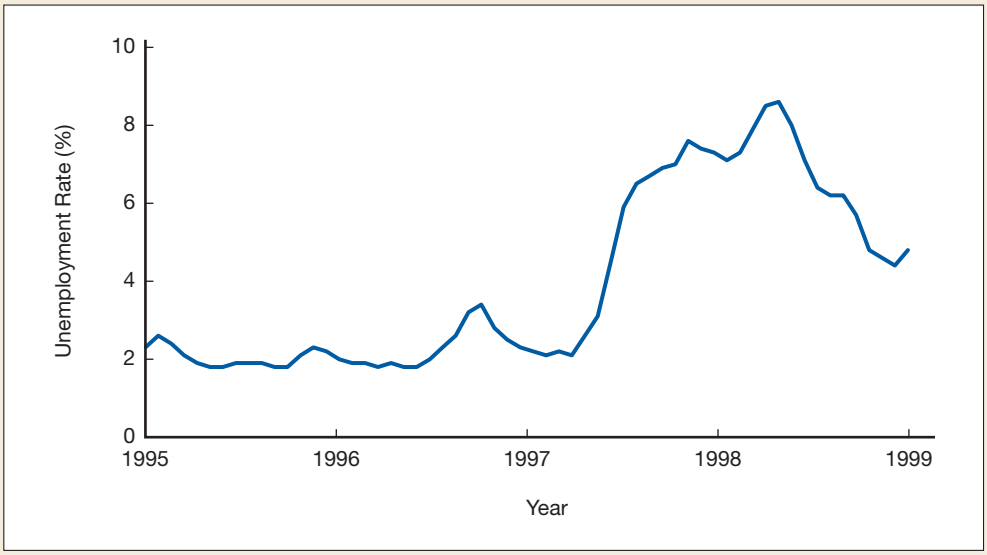
Starting in the early 1990s, the South Korean government removed many restrictive regulations on financial institutions in an effort to liberalize the country's financial markets and also embarked on a financial globalization process by opening up South Korea's capital markets to capital flows from abroad. This resulted in a lending boom during which bank credit to the private nonfinancial business sector accelerated

FIGURE S1-4

Unemployment, South Korea, 1995–1999

The unemployment rate was below 3% before the crisis but jumped to above 8% during the crisis.

Source: International Monetary Fund. *International Financial Statistics.* www.imfstatistics.org/imf/.



sharply, with lending fuelled by massive foreign borrowing that expanded at rates close to 20% per year. Because of weak bank regulator supervision and a lack of expertise in screening and monitoring borrowers at banking institutions, losses on loans began to mount, causing an erosion of banks' net worth (capital).

Perversion of the Financial Liberalization and Globalization Process: Chaebols and the South Korean Crisis

Powerful business interests play an active role in the mismanagement of financial liberalization. Nowhere was this clearer than in South Korea before the financial crisis of the late 1990s, when large family-owned conglomerates known as chaebols dominated the economy, with sales amounting to nearly 50% of the country's GDP.

The chaebols were politically very powerful and deemed "too big to fail" by the government. With this implicit guarantee, the chaebols knew they would receive direct government assistance or directed credit if they got into trouble, but they could keep all of the profits if their bets paid off. Not surprisingly, chaebols borrowed like crazy and became highly leveraged.

In the 1990s, the chaebols weren't making any money. From 1993 to 1996, the return on assets for the top 30 chaebols was never much more than 3% (a comparable figure for U.S. corporations is 15%–20%). In 1996, right before the crisis hit, the average rate of return on assets had fallen to 0.2%. Furthermore, only the top five most profitable chaebols had any profits, while the chaebols in places 6 to 30 (in terms of profitability) never had a rate of return on assets much above 1% and in many years had negative rates of return. With this poor profitability record and the already high leverage, any banker would pull back on lending to these conglomerates *if* there were no government safety net. Because the banks knew the government would make good on the chaebols' loans if they were in default, banks continued to lend to the chaebols. This substantial financing from commercial banks, however, was not enough to feed the chaebols' insatiable appetite for more credit. The chaebols decided that the way out of their troubles was to pursue growth, and they needed massive amounts of funds to do so. Even with the vaunted South Korean national savings rate of over 30%, there just weren't enough loanable funds available to finance the chaebols' planned expansion. Where could they get the funds they desired? The answer was in the international capital markets.

The chaebols encouraged the South Korean government to accelerate the process of opening up South Korean financial markets to foreign capital as part of the liberalization process. In 1993, the government expanded the ability of domestic banks to make loans denominated in foreign currency. At the same time, the South Korean government effectively allowed unlimited short-term foreign borrowing by financial institutions but maintained quantity restrictions on long-term borrowing as a means of managing capital inflows into the country. Opening up in the short term but not in the long term to foreign capital inflows made no economic sense. Short-term capital flows make an emerging-market economy financially fragile: Short-term capital can fly out of the country extremely rapidly if there is any whiff of a crisis.

Opening up primarily to short-term capital, however, made complete political sense: The chaebols needed the money, and it is much easier to borrow short-term funds at low interest rates in the international market than to borrow long-term funds, because long-term lending is much riskier for foreign creditors. In the aftermath of the government's new banking policies, South Korean banks opened 28 branches in foreign countries that gave them access to foreign funds.

Although South Korean financial institutions now had access to foreign capital, the chaebols still had a problem. Because they were not allowed to own commercial banks, the chaebols might not get all the bank loans they needed. However, a type of financial institution specific to South Korea existed that *would* enable them to get the loans they needed: the merchant bank. Merchant banking corporations were wholesale financial institutions that engaged in underwriting securities, leasing, and short-term lending to the corporate sector. They obtained funds for these loans by issuing bonds and commercial paper and by borrowing from interbank and foreign markets. At the time of the Korean crisis, merchant banks were allowed to borrow abroad and were virtually unregulated. The chaebols saw their opportunity and convinced government officials to convert many finance companies (some already owned by the chaebols) into merchant banks. The merchant banks channelled massive amounts of funds to their chaebol owners, where they flowed into unproductive investments in steel, automobile production, and chemicals. When the loans went sour, the stage was set for a disastrous financial crisis.

Stock Market Decline and Failure of Firms Increase Uncertainty

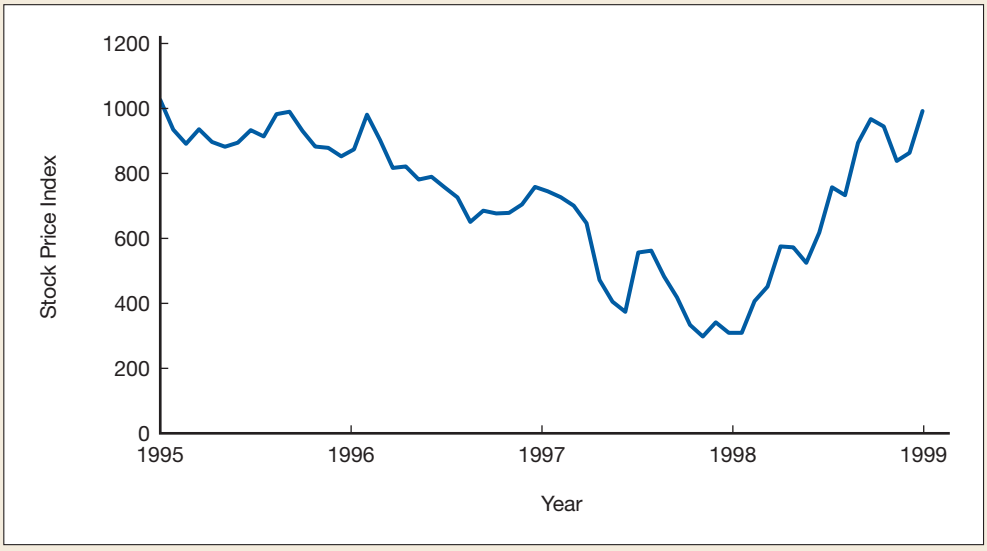
The South Korean economy then experienced a negative shock to export prices that hurt the chaebols' already-thin profit margins and the small- and medium-sized firms that were tied to them. On January 23, 1997, a second major shock occurred, creating great uncertainty for the financial system: Hanbo, the 14th largest chaebol, declared bankruptcy. Indeed, the bankruptcy of Hanbo was just the beginning. Five more of the 30 largest chaebols declared bankruptcy before the year was over. The greater uncertainty created by these bankruptcies and the deteriorating condition of financial and nonfinancial balance sheets caused the stock market to decline sharply, by more than 50% from its peak, as shown in Figure S1-5.

FIGURE S1-5

Stock Market Index, South Korea, 1995–1999

Stock prices fell by over 50% during the crisis.

Source: Global Financial Data, available at www.globalfinancialdata.com/.



Adverse Selection and Moral Hazard Problems Worsen, and the Economy Contracts

As we have seen, an increase in uncertainty and a decrease in net worth resulting from a stock market decline will exacerbate asymmetric information problems. It becomes hard to screen out good borrowers from bad ones. The decline in net worth decreases the value of firms' collateral and increases their incentives to make risky investments because there is less equity to lose if the investments are unsuccessful. The increase in uncertainty and stock market declines that occurred before the South Korean crisis, along with the deterioration in banks' balance sheets, worsened adverse selection and moral hazard problems. As a result, lending declined and the economy weakened. The weakening of the economy, along with the deterioration of bank balance sheets, ripened the South Korean economy for the next stage, a currency crisis that would send the economy into a full-fledged financial crisis and a depression.

Currency Crisis Ensues

Given the weakness of balance sheets in the financial sector and the increased exposure of the economy to sudden stops in capital inflows caused by the large amount of short-term external borrowing, a speculative attack on South Korea's currency was inevitable. With the collapse of the Thai currency, the baht, in July 1997 and the announced closing of 42 finance companies in Thailand in early August 1997, contagion began to spread as participants in the market wondered whether similar problems existed in other East Asian countries. Soon speculators recognized that the banking sector in South Korea was in trouble. They knew the South Korean central bank could no longer defend the currency by raising interest rates because this action would sink the already weakened banks. Speculators pulled out of the South Korean currency, the won, leading to a speculative attack.

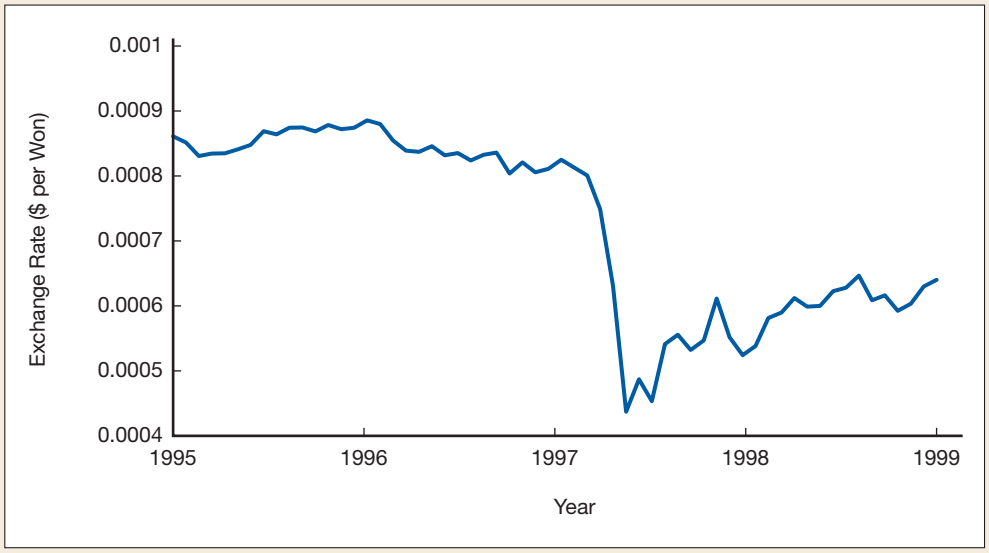
Final Stage: Currency Crisis Triggers Full-Fledged Financial Crisis

The speculative attack led to a sharp drop in the value of the won, by nearly 50%, as shown in Figure S1-6. Because both nonfinancial and financial firms had so much foreign-currency debt, the nearly 50% depreciation of the Korean won doubled the value of the foreign-denominated debt in terms of the domestic currency and therefore led to a severe erosion of net worth. This loss of net worth led to a severe increase in adverse selection and moral hazard problems in South Korean financial markets, for domestic and foreign lenders alike.

The deterioration in firms' cash flow and balance sheets worsened the banking crisis. Bank balance sheets were devastated when the banks paid off their foreign-currency borrowing with more Korean won and yet could not collect on the dollar-denominated loans they had made to domestic firms. In addition, the fact that financial institutions had been encouraged to make their foreign borrowing short term increased their liquidity problems because the banks had to pay these loans back so quickly. The government stepped in to guarantee all bank deposits and prevent a bank panic, but the loss of capital meant that banks had to curtail their lending.

In 1998, real GDP fell at over a 6% rate, as shown in Figure S1-3, and unemployment rose sharply (refer to Figure S1-4). In this situation, we might expect the inflation rate to fall as well. Inflation, however, did not fall; instead it rose, as shown in Figure S1-2. The currency crisis was behind this key difference. Specifically, the collapse of the

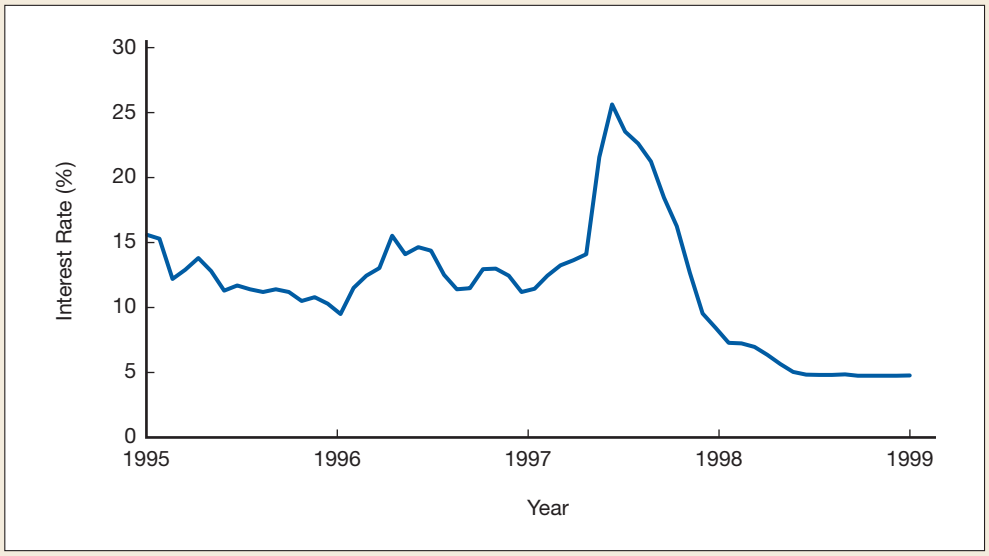
FIGURE S1-6
Value of South Korean Currency, 1995–1999
The Korean won lost nearly half its value during the crisis.
Source: International Monetary Fund. *International Financial Statistics.* www.imfstatistics.org/imf/.



South Korean currency after the successful speculative attack on it raised import prices, which fed directly into inflation and weakened the credibility of the Bank of Korea as an inflation fighter. The rise in import prices led to a price shock, while the weakened credibility of the Bank of Korea led to a rise in expected inflation, with inflation climbing sharply from around the 5% level to nearly 10%.

Market interest rates soared to over 20% by the end of 1997 (Figure S1-7) to compensate for the high inflation. They also rose because the Bank of Korea pursued a tight monetary policy in line with recommendations from the International Monetary Fund.

FIGURE S1-7
Interest Rates, South Korea, 1995–1999
Market interest rates soared to 25% during the crisis.
Source: International Monetary Fund. *International Financial Statistics.* www.imfstatistics.org/imf/.



High interest rates led to a drop in cash flows, which forced firms to obtain external funds and increased adverse selection and moral hazard problems in the credit markets. The increase in asymmetric information problems in the credit markets, along with the direct effect of higher interest rates on investment decisions, led to a further contraction in investment spending, providing another reason for the falling economic activity.

Recovery Commences

In 1998, the South Korean government responded very aggressively to the crisis by implementing a series of financial reforms that helped restore confidence in the financial system. Financial markets began to recover, which helped stimulate lending, and the economy finally began to recover.

The South Korean financial crisis inflicted a high human cost, too. The ranks of the poor swelled from 6 million to over 10 million, suicide and divorce rates jumped by nearly 50%, drug addiction rates climbed by 35%, and the crime rate rose by over 15%.

APPLICATION The Argentine Financial Crisis, 2001–2002

Argentina's financial crisis was triggered by severe fiscal imbalances. We will now examine the details of the crisis.

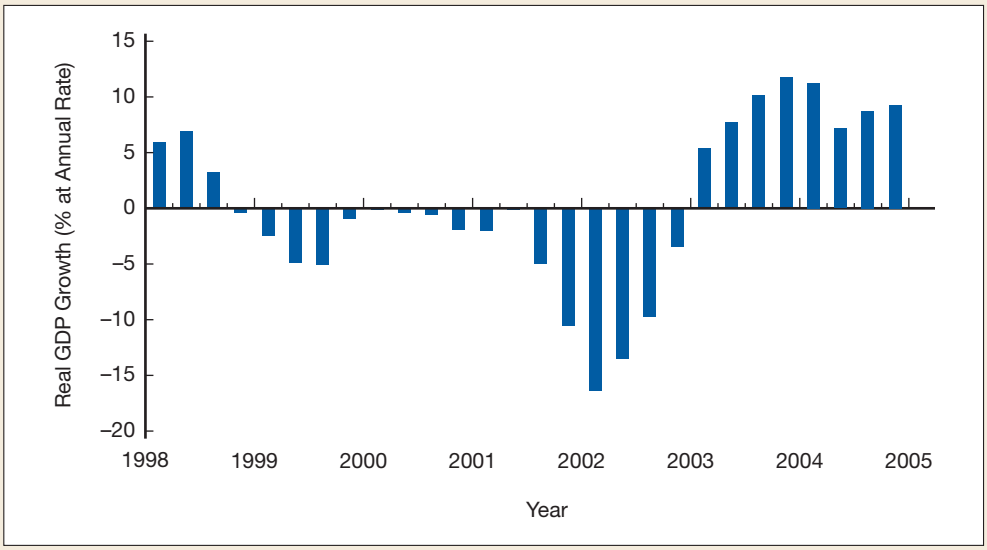
Severe Fiscal Imbalances

In contrast to Mexico and the East Asian countries, Argentina had a well-supervised banking system, and a lending boom did not occur before Argentina's crisis. The banks were in surprisingly good shape before the crisis, even though a severe recession had begun in 1998 (see Figure S1-8). Unfortunately, however, Argentina has always had

FIGURE S1-8
Real GDP Growth, Argentina, 1998–2004

Real GDP growth collapsed during the crisis, declining at an annual rate of over 15%.

Source: International Monetary Fund. *International Financial Statistics.* www.imfstatistics.org/imf/.



difficulty controlling its budget. In Argentina, the provinces (similar to provinces in Canada and states in the United States) control a large percentage of public spending, but the responsibility for raising revenue is left primarily to the federal government. With this system, the provinces have incentives to spend beyond their means and then call on the federal government periodically to assume responsibility for their debt. As a result, Argentina's government budget is perennially in deficit.

The recession starting in 1998 made the situation even worse because it led to declining tax revenues and a widening gap between government expenditures and taxes. The subsequent severe fiscal imbalances were so large that the government had trouble getting both domestic residents and foreigners to buy enough of its bonds, so it coerced banks into absorbing large amounts of government debt. By 2001, investors were losing confidence in the ability of the Argentine government to repay this debt. The value of the debt plummeted, leaving big holes in banks' balance sheets. What had once been considered one of the best-supervised and strongest banking systems among emerging-market countries was now losing deposits.

Adverse Selection and Moral Hazard Problems Worsen

The deterioration of bank balance sheets and the loss of deposits led the banks to cut back on their lending. As a result, adverse selection and moral hazard problems worsened and lending began to decline. The resulting weakening of the economy and deterioration of bank balance sheets set the stage for the next stage of the crisis, a bank panic.

Bank Panic Begins

In October 2001, negotiations between the central government and the provinces to improve the fiscal situation broke down, and tax revenues continued to fall as the economy suffered. Default on government bonds was now inevitable. As a result, a full-fledged bank panic began in November, with deposit outflows running nearly $1 billion a day. At the beginning of December, the government was forced to close the banks temporarily and impose a restriction called the *corralito* (small fence), under which depositors could withdraw only 250 pesos ($250) in cash per week. The corralito was particularly devastating for the poor, who were highly dependent on cash to conduct their daily transactions. The social fabric of Argentine society began to unravel. Nearly 30 people died in violent riots.

Currency Crisis Ensues

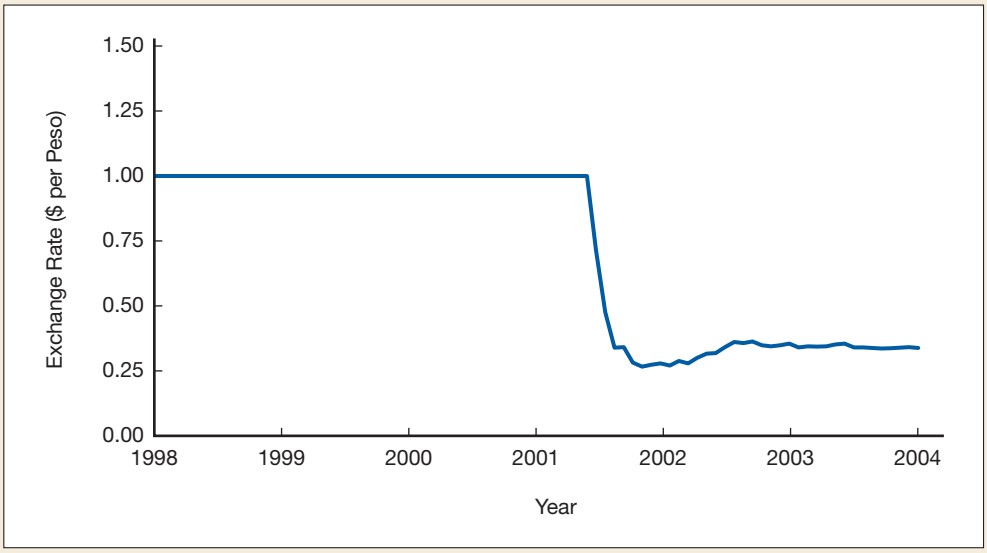
The bank panic signalled that the government could no longer allow interest rates to remain high in order to prop up the value of the peso and preserve the currency board, an arrangement in which the Argentine government fixed the value of one Argentine peso to one U.S. dollar by agreeing to buy and sell pesos at that exchange rate (discussed in Chapter 19). Raising interest rates to preserve the currency board was no longer an option because it would have meant destroying the already weakened banks. The public now recognized that the peso would have to decline in value in the near future, so a speculative attack began in which people sold pesos for dollars. In addition, the government's dire fiscal position made it unable to pay back its debt, providing yet another reason for the investors to pull money out of the country, leading to further peso sales.

FIGURE S1-9

Argentine Peso, 1998–2004

The value of the Argentine peso fell from US$1.00 before the crisis to under US$0.30 by June 2002.

Source: International Monetary Fund. *International Financial Statistics.* www.imfstatistics.org/imf/.



On December 23, 2001, the government announced the inevitable: a suspension of external debt payments for at least 60 days. Then, on January 2, 2002, the government abandoned the currency board.

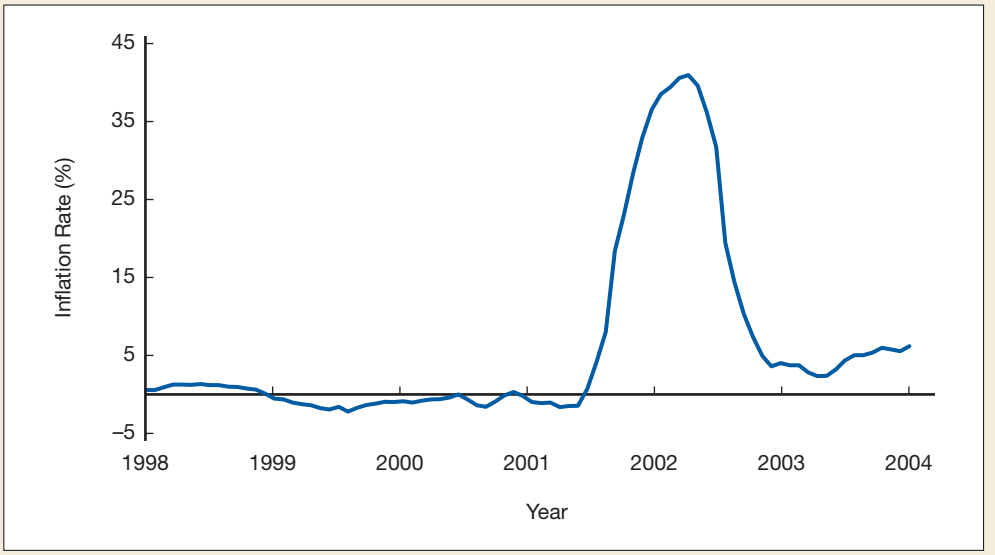
Currency Crisis Triggers Full-Fledged Financial Crisis

The peso now went into free fall, dropping from a value of US$1.00 to less than US$0.30 by June 2002 and then stabilizing at around US$0.33 thereafter, as shown in Figure S1-9. Because Argentina had a higher percentage of debt denominated in dollars than any of the other crisis countries, the effects of the peso collapse on balance sheets were particularly devastating. With the peso falling to one-third of its value before the crisis, all dollar-denominated debt tripled in peso terms. Since Argentina's tradable sector was small, most businesses' production was priced in pesos. If they had to pay back their dollar debt, almost all firms would become insolvent. In this environment, financial markets could not function because net worth would not be available to mitigate adverse selection and moral hazard problems.

Given the losses on the defaulted government debt and the rising loan losses, Argentine banks found their balance sheets in a precarious state. Further, the run on the banks had led to huge deposit outflows. Lacking the resources necessary to make new loans, the banks could no longer solve adverse selection and moral hazard problems. The government bond default and the conditions of the Argentine financial markets meant that foreigners were unwilling to lend and instead were actually pulling their money out of the country.

With the financial system in jeopardy, financial flows came to a grinding halt. The resulting curtailment of lending led to a further contraction of economic activity. The corralito also played an important role in the weakening of the economy. By making it more difficult to get cash, the corralito caused a sharp slowdown in the underground economy, which in Argentina is large and runs primarily on cash.

FIGURE S1-10
Inflation, Argentina, 1998–2004
Inflation surged to over 40% during the crisis.
Source: International Monetary Fund. *International Financial Statistics.* www.imfstatistics.org/imf/.



Just as in South Korea, the collapse of the Argentine currency after the successful speculative attack on the currency raised import prices, which fed directly into inflation and weakened the credibility of the Argentine central bank to keep inflation under control. Indeed, Argentina's history of very high inflation meant that there was an even larger rise in expected inflation in Argentina than in South Korea. Inflation in Argentina rose as high as 40% at an annual rate, as shown in Figure S1-10. Because the rise in actual inflation was accompanied by a rise in expected inflation, interest rates went to even higher levels, as indicated in Figure S1-11. The

FIGURE S1-11
Interest Rates, Argentina, 1998–2004
Interest rates rose to nearly 100% during the crisis.
Source: International Monetary Fund. *International Financial Statistics.* www.imfstatistics.org/imf/.

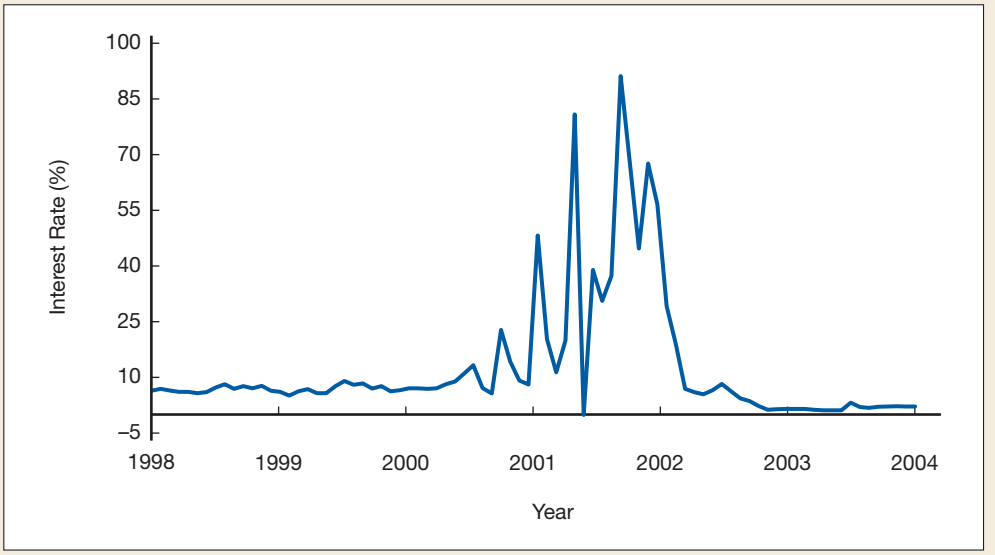
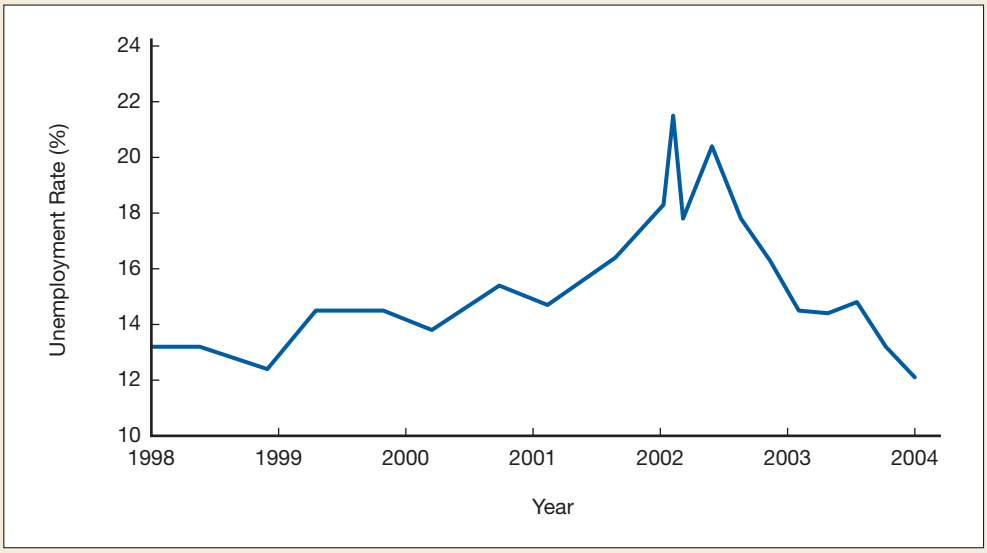


FIGURE S1-12

Unemployment Rate, Argentina, 1998–2004

The unemployment rate climbed to over 20% during the crisis.

Source: Global Financial Data, available at www.globalfinancialdata.com/.



higher interest payments led to a decline in the cash flow of both households and businesses, which now had to seek external funds to finance their investments. Given the uncertainty in financial markets, asymmetric information problems were particularly severe, and this meant that investment could not be funded. Households and businesses cut back their spending further. As predicted by our aggregate demand and supply analysis, the Argentine economy plummeted. In the first quarter of 2002, output was falling at an annual rate of more than 15%, as shown in Figure S1-8. Figure S1-12 demonstrates that unemployment shot up to over 20%. The increase in poverty was dramatic: The percentage of the Argentine population living in poverty rose to almost 50% in 2002. Argentina was experiencing the worst depression in its history—one every bit as bad as, and maybe even worse than, the U.S. Great Depression.

Recovery Begins

As the financial crisis receded, a boom in the demand for Argentina's commodity exports encouraged the beginnings of recovery in the Argentine economy. By the end of 2003, economic growth was running at an annual rate of around 10%, and as Figure S1-12 demonstrates, unemployment had fallen to below 15%. Inflation also had fallen, to below 5%, as shown in Figure S1-10.

Although we have drawn a strong distinction between financial crises in emerging-market economies and those in advanced economies, there have been financial crises in advanced economies that have had much in common with financial crises in emerging-market economies. This is illustrated by the Global box "When an Advanced Economy Is Like an Emerging-Market Economy: The Icelandic Financial Crisis of 2008." ◆

Global **When an Advanced Economy Is Like an Emerging-Market Economy: The Icelandic Financial Crisis of 2008**

The financial crisis and economic contraction in Iceland that started in 2008 followed the script of a financial crisis in an emerging-market economy, a remarkable turn of events for a small, wealthy nation with one of the world's highest standards of living.

As part of a financial liberalization process, in 2003 the government of Iceland sold its state-owned banks to local investors who supersized the Icelandic banking industry. These investors set up overseas branches to take foreign-currency deposits from thousands of Dutch and British households and borrowed heavily from short-term wholesale funding markets in which foreign-denominated credit was ample and cheap. The investors channelled the funds to local investment firms, many of which had ties to the bank owners themselves. Not surprisingly, many of the funds went into high-risk investments like equities and real estate. Iceland's stock market value ballooned to 250% of GDP, and firms, households, and banks borrowed heavily in foreign currencies, leading to a severe currency mismatch such as had occurred in many emerging-market countries. Meanwhile, Iceland's regulatory system provided ineffective supervision of bank risk-taking.

In October 2008, the failure of the U.S. investment bank Lehman Brothers shut down the wholesale lending system that had kept the Icelandic banks afloat. Without access to funding, the banks couldn't repay their foreign-currency debts, and much of the collateral they held—including shares in other Icelandic banks—fell to a fraction of its previous value. Banks had grown so large—their assets were nearly nine times the nation's gross domestic product—that not even the government could credibly rescue the banks from failure. Foreign capital fled Iceland, and the value of the Icelandic krona tumbled by over 50%.

As a result of the currency collapse, the debt burden in terms of domestic currency more than doubled, destroying a large part of Icelandic firms' and households' net worth, leading to a full-scale financial crisis. The economy went into a severe recession, with the unemployment rate tripling, real wages falling, and the government incurring huge budget deficits. Relationships with foreign creditors became tense, and the United Kingdom even froze assets of Icelandic firms under an antiterrorism law. Many Icelandic people returned to traditional jobs in the fishing and agricultural industries, where citizens could find a silver lining: The collapse of the Icelandic currency made exports of Iceland's famous cod fish, Arctic char, and Atlantic salmon more affordable to foreigners.

S1.2 PREVENTING EMERGING-MARKET FINANCIAL CRISES

LO S1.2 Discuss the different measures used to prevent emerging-market financial crises.

Our analysis of financial crises in emerging-market economies suggests a number of government policies that could help make financial crises in emerging-market countries less likely.[2]

Beef Up Prudential Regulation and Supervision of Banks

As we have seen, the banking sector sits at the root of financial crises in emerging-market economies. To prevent crises, therefore, governments must improve prudential regulation and supervision of banks to limit their risk-taking. First, regulators should ensure that banks hold ample capital to cushion losses from economic shocks and to give bank owners, who have more to lose, an incentive to pursue safer investments.

[2] For a more extensive discussion of reforms to prevent financial crises in emerging-market countries, see Chapter 9, "Preventing Financial Crises," in Frederic S. Mishkin, *The Next Great Globalization: How Disadvantaged Nations Can Harness Their Financial Systems to Get Rich* (Princeton, NJ: Princeton University Press, 2006), 137–163.

Prudential supervision can also help promote a safer and sounder banking system by ensuring that banks have proper risk management procedures in place, including (1) good risk measurement and monitoring systems, (2) policies to limit activities that present significant risks, and (3) internal controls to prevent fraud or unauthorized activities by employees. As indicated in the South Korea example, regulations should also ban commercial businesses from owning banking institutions. When commercial businesses own banks, they are likely to use them to channel lending to themselves, as the chaebols did in South Korea, leading to risky lending that can provoke a banking crisis.

For prudential supervision to work, prudential supervisors must have adequate resources with which to do their jobs. This is a particularly serious problem in emerging-market countries, where prudential supervisors earn low salaries and lack basic tools such as computers. Because politicians often put pressure on prudential supervisors to discourage them from being "too tough" on banks that make political contributions (or outright bribes), a more independent regulatory and supervisory agency can better withstand political influence, increasing the likelihood that prudential supervisors will do their jobs and limit bank risk-taking.

Encourage Disclosure and Market-Based Discipline

The public sector, acting through prudential regulators and supervisors, will always struggle to control risk-taking by financial institutions. Financial institutions have incentives to hide information from bank supervisors in order to avoid restrictions on their activities, and they can become quite adept and crafty at masking risk. Also, supervisors may be corrupt or may give in to political pressure and so not do their jobs properly.

To eliminate these problems, financial markets need to discipline financial institutions from taking on too much risk. Government regulations to promote disclosure by banking and other financial institutions of their balance sheet positions, therefore, are needed to encourage these institutions to hold more capital, because depositors and creditors will be unwilling to put their money into an institution that is thinly capitalized. Regulations to promote disclosure of banks' activities will also encourage depositors and creditors to pull their money out of institutions that are engaged in risky activities and put it into institutions that are behaving responsibly.

Limit Currency Mismatch

As we have seen, emerging-market financial systems can be very vulnerable to declines in the value of the nation's currency. Often, firms in these countries borrow in foreign currencies even though their products and assets are priced in the domestic currency. A collapse of the domestic currency causes debt denominated in foreign currencies to become particularly burdensome because it has to be paid back in the more expensive foreign currency, thereby causing a deterioration in firms' balance sheets. Governments can limit currency mismatch by implementing regulations or taxes that discourage the issuance by nonfinancial firms of debt denominated in foreign currencies. Regulation of banks also can limit bank borrowing in foreign currencies.

Moving to a flexible exchange rate regime, in which exchange rates fluctuate, can help discourage borrowing in foreign currencies because there is now more risk associated with doing so. Monetary policy that promotes price stability also helps by making the domestic currency less subject to decreases in its value as a result of high inflation, thus making it more desirable for firms to borrow in domestic rather than foreign currency.

Sequence Financial Liberalization

Although financial liberalization can be highly beneficial in the long run, our analysis of financial crises in emerging-market economies in this chapter shows that if this process is not managed properly, it can be disastrous. If the proper bank regulatory/supervisory structure and disclosure requirements are not in place when liberalization occurs, the necessary constraints on risk-taking behaviour will be far too weak.

To avoid financial crises, policymakers need to put in place the proper institutional infrastructure before liberalizing their financial systems. Crucial to avoiding financial crises is the implementation of the policies described previously, which involve strong prudential regulation and supervision and limits on currency mismatch. Because implementing these policies takes time, financial liberalization may have to be phased in gradually, with some restrictions on credit issuance imposed along the way.

SUMMARY

1. Financial crises in emerging-market countries develop along two basic paths: one involving the mismanagement of financial liberalization/globalization that weakens bank balance sheets and the other involving severe fiscal imbalances. Both lead to a speculative attack on the currency and eventually to a currency crisis in which there is a sharp decline in the value of the domestic currency. The decline in the value of the domestic currency causes a sharp rise in the debt burden of domestic firms, which leads to a decline in firms' net worth as well as increases in inflation and interest rates. Adverse selection and moral hazard problems then worsen, leading to a collapse of lending and economic activity. The worsening economic conditions and increases in interest rates result in substantial losses for banks, leading to a banking crisis, which further depresses lending and aggregate economic activity.
2. The financial crisis in South Korea followed the pattern described above. It started with a mismanagement of financial liberalization and globalization, followed by a stock market crash and a failure of firms that increased uncertainty, a worsening of adverse selection problems, a currency crisis, and then a full-fledged financial crisis. The financial crisis led to a sharp contraction in economic activity and a rise in inflation as well as a weakening of the social fabric.
3. In contrast to the crisis in South Korea, the Argentine financial crisis started with severe fiscal imbalances. A recession that had been going on since 1998, along with a deterioration of bank balance sheets that occurred when fiscal imbalances led to losses from government bonds that the banks had on their balance sheets, led to a worsening of adverse selection and moral hazard problems, a bank panic, a currency crisis, and then a full-fledged financial crisis. As in South Korea, the financial crisis in Argentina led to a decline in economic activity and a rise in inflation, but the economic contraction and rise in inflation were even worse because Argentina's central bank lacked credibility as an inflation fighter.
4. Policies to prevent financial crises in emerging-market economies include improving prudential regulation and supervision, limiting currency mismatch, and sequencing of financial liberalization.

MyLab Economics Make the grade with MyLab Economics: Select end-of-chapter questions can be found on MyLab Economics. You can practise them as often as you want, and most feature step-by-step guided instructions to help you find the right answer.

QUESTIONS

1. What are the two basic causes of financial crises in emerging-market economies?
2. Why might financial liberalization and globalization lead to financial crises in emerging-market economies?
3. Why might severe fiscal imbalances lead to financial crises in emerging-market economies?
4. What other factors can initiate financial crises in emerging-market economies?
5. What events can ignite a currency crisis?
6. Why do currency crises make financial crises in emerging-market economies even more severe?
7. How did the financial crises in South Korea and Argentina affect aggregate demand, short-run aggregate supply, and output and inflation in these countries?
8. What can emerging-market countries do to strengthen prudential regulation and supervision of their banking systems? How might these steps help avoid future financial crises?
9. How can emerging-market economies avoid the problems of currency mismatch?
10. Why might emerging-market economies want to implement financial liberalization and globalization gradually rather than all at once?

PART 1

Introduction

Crisis and Response: The Global Financial and the Covid-19 Crises

In August 2007, financial markets began to seize up, and over the next two years the world economy experienced a global financial crisis that was the most severe since the Great Depression years of the 1930s. The financial crisis started in the United States and was the result of a credit-driven, asset price bubble in the U.S. housing market. When that bubble burst, housing prices in the United States plummeted, the stock market crashed, unemployment skyrocketed, and both businesses and households found they couldn't get credit. Not only did the central bank of the United States, the Federal Reserve, respond by sharply lowering interest rates and intervening in credit markets to provide them with massive amounts of liquidity, but central banks and governments around the world worked on full-scale banking bailouts and rescue packages adding up to trillions of dollars. However, even with these aggressive actions aimed at stabilizing the financial system and boosting their economies, it took 10 years before economies around the world returned to full employment.

In 2020, the world economy was hit by another crisis, but this time the source was not man-made but from a strain of coronavirus that originated in Wuhan, China. On March 11, 2020, the World Health Organization declared a world pandemic, and economies all over the world began to lock down and the stock market crashed. Central banks stepped in again to lower interest rates and provide massive amounts of liquidity to shore up the financial system and stimulate lending by banks and other institutions. Governments passed many fiscal stimulus packages, totalling, for example, close to $250 billion in Canada and over US$3 trillion in the United States, the largest stimulus packages in history. Despite these efforts, by April 2020, unemployment rates had risen to close to 15%, the highest levels since the Great Depression of the 1930s.

The aftermath of the global financial and the coronavirus crises demonstrates the importance of banks and financial systems to economic well-being as well as the major role of money in the economy. Part 1 of this book introduces the study of money, banking, and financial markets. Chapter 1 outlines a road map of the book and discusses why it is so worthwhile to study money, banking, and financial markets. Chapter 2 provides a general overview of the financial system. Chapter 3 then explains what money is and how it is measured.

1 Why Study Money, Banking, and Financial Markets?

Learning Objectives

1.1 Recognize the importance of financial markets in the economy.

1.2 Describe how financial intermediation and financial innovation affect banking and the economy.

1.3 Identify the basic links among monetary policy, the business cycle, and economic variables.

1.4 Explain the importance of exchange rates in a global economy.

1.5 Explain how the study of money, banking, and financial markets may advance your career.

1.6 Describe how the text approaches the teaching of money, banking, and financial markets.

Preview

You have just heard on the evening news that the Bank of Canada is raising the overnight interest rate by $\frac{1}{2}$ of a percentage point. What effect might this have on the interest rate of an automobile loan when you finance your purchase of a sleek new sports car? Does it mean that a house will be more or less affordable in the future? Will it make it easier or harder for you to get a job next year?

This book provides answers to these and other questions by examining how financial markets (such as those for bonds, stocks, and foreign exchange) and financial institutions (banks, insurance companies, mutual funds, and other institutions) work and by exploring the role of money in the economy. Financial markets and institutions affect not only your everyday life but also the flow of trillions of dollars of funds throughout our economy, which in turn affects business profits, the production of goods and services, and even the economic well-being of countries other than Canada. What happens to financial markets, financial institutions, and money is of great concern to politicians and can have a major impact on elections. The study of money, banking, and financial markets will reward you with an understanding of many exciting issues. In this chapter, we provide a road map of this book by outlining these issues and exploring why they are worth studying.

1.1 WHY STUDY FINANCIAL MARKETS?

LO 1.1 Recognize the importance of financial markets in the economy.

Part 2 of this book focuses on **financial markets**—markets in which funds are transferred from people who have an excess of available funds to people who have a shortage. Financial markets, such as bond and stock markets, are crucial to promoting greater economic efficiency by channelling funds from people who do not have a productive use for them to those who do. Indeed, well-functioning financial markets are a key factor in producing high economic growth, and poorly performing financial markets are one reason that many countries in the world remain desperately poor. Activities in financial markets also have a direct effect on personal wealth, the behaviour of businesses and consumers, and the cyclical performance of the economy.

Debt Markets and Interest Rates

A **security** (also called a *financial instrument*) is a claim on the issuer's future income or **assets** (any financial claim or piece of property that is subject to ownership). A **bond** is a debt security that promises to make periodic payments for a specified period of time.[1] Debt markets, also often generically referred to as *bond markets*, are especially important to economic activity because they enable corporations and governments to borrow money to finance their activities, and because it is where interest rates are determined. An **interest rate** is the cost of borrowing or the price paid for the rental of funds (usually expressed as a percentage of the rental of $100 per year). Many types of interest rates are found in the economy—mortgage interest rates, car loan rates, and interest rates on many different types of bonds.

Interest rates are important on a number of levels. On a personal level, high interest rates might deter you from buying a house or a car because the cost of financing would be high. Conversely, high interest rates might encourage you to save because you can earn more interest income by putting aside some of your earnings as savings. On a more general level, interest rates have an impact on the overall health of the economy because they affect not only consumers' willingness to spend or save but also businesses' investment decisions. High interest rates, for example, might cause a corporation to postpone building a new plant that would provide more jobs.

Because changes in interest rates affect individuals, financial institutions, businesses, and the overall economy, it is important to explain substantial fluctuations in interest rates over the past 40 years. For example, the interest rate on three-month Treasury bills peaked at over 20% in August 1981. This interest rate then fell to a low of 2% in the early 2000s and rose to above 4% by 2007. It then fell to near zero from 2009 to 2015, then rose to close to 2% in 2018, only to fall back to near zero again when the coronavirus pandemic led to the Covid-19 recession in March 2020.

Because different interest rates have a tendency to move in unison, economists frequently lump interest rates together and refer to "the" interest rate. As Figure 1-1 shows, however, interest rates on several types of bonds can differ substantially. The interest rate on three-month Treasury bills, for example, fluctuates more than the other interest rates and is lower on average. The interest rate on long-term corporate bonds is higher, on average, than the other interest rates, and the spread between it and the other rates fluctuates over time. It became larger in the 1970s, narrowed in the 1990s, rose briefly in the early 2000s, narrowed again, and then rose sharply starting in the summer of 2007. It then began to decline toward the end of 2009, returning to low levels by 2018, and then rose again during the Covid-19 recession in 2020.

In Chapter 2 we study the role of bond markets in the economy, and in Chapters 4 through 6 we examine what an interest rate is, how the common movements in interest rates come about, and why the interest rates on different bonds vary.

The Stock Market

A **common stock** (typically called simply a **stock**) represents a share of ownership in a corporation. It is a security that is a claim on the earnings and assets of the

[1]The definition of *bond* used throughout this book is the broad one commonly used in academic settings, which covers both short- and long-term debt instruments. However, some practitioners in financial markets use the word *bond* to describe only specific long-term debt instruments such as corporate bonds or Government of Canada bonds.

FIGURE 1-1 Interest Rates on Selected Bonds, 1977–2020

Although different interest rates have a tendency to move in unison, they often differ substantially, and the spreads between them fluctuate.

Source: Statistics Canada CANSIM series V122531, V122544, and V122518, the latter extended with the 30-year A-rated corporate bond interest rate (C28730Y) from Bloomberg.

corporation. Issuing stock and selling it to the public is a way for corporations to raise funds to finance their activities. The stock market, in which claims on the earnings of corporations (shares of stock) are traded, is the most widely followed financial market in almost every country that has one; that's why it's often called simply "the market." A big swing in the prices of shares in the stock market is always a major story on the evening news. People often speculate on where the market is heading and get very excited when they can brag about their latest "big killing," but they become depressed when they suffer a big loss. The attention the market receives can probably be best explained by one simple fact: It is a place where people can get rich—or poor—very quickly.

As Figure 1-2 indicates, stock prices are extremely volatile. After rising steadily during the 1980s, the market experienced the worst one-day drop in its entire history on October 19, 1987—"Black Monday"—when the S&P/TSX Composite Index fell by 11%. From then until 2000, the stock market experienced one of the greatest rises (often referred to as a "bull market") in its history, with the S&P/TSX climbing to a peak of over 11 000. With the collapse of the high-tech bubble in 2000, the stock market fell sharply, dropping by over 40% by late 2002. It then rose to an all-time high above the 14 000 level in early 2008, only to fall by over 50% of its value to a low of close to 8000 in early 2009. Another bull market then began, with the S&P/TSX reaching new highs above 17 000 by 2020. These considerable fluctuations in stock prices affect the size of people's wealth and, as a result, their willingness to spend.

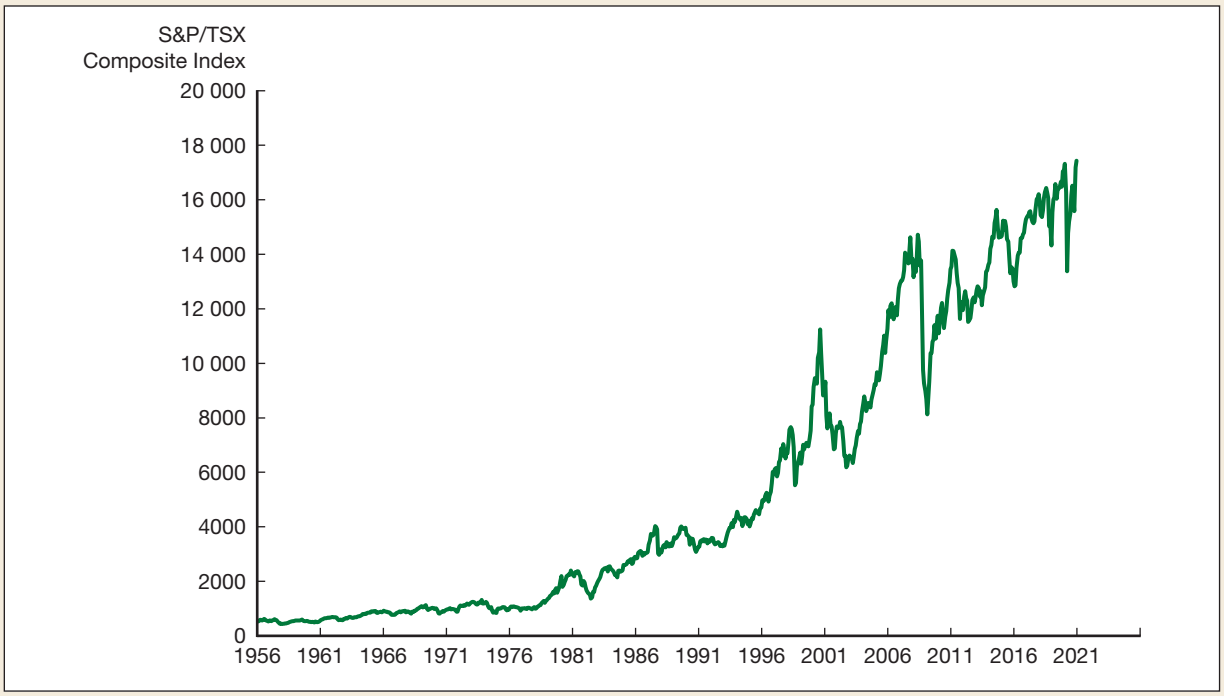


FIGURE 1-2 **Stock Prices as Measured by the S&P/TSX Composite Index, 1956–2020**
Stock prices are extremely volatile.
Source: Statistics Canada CANSIM series V122620.

The stock market is also an important factor in business investment decisions, because the price of shares affects the amount of funds that can be raised by selling newly issued stock to finance investment spending. A higher price for a firm's shares means that the firm can raise a larger amount of funds, which it can then use to buy production facilities and equipment.

In Chapter 2 we examine the role the stock market plays in the financial system, and in Chapter 7 we return to the issue of how stock prices behave and respond to information in the marketplace.

1.2 WHY STUDY FINANCIAL INSTITUTIONS AND BANKING?

LO 1.2 Describe how financial intermediation and financial innovation affect banking and the economy.

Part 3 of this book focuses on financial institutions and the business of banking. Banks and other financial institutions are what make financial markets work. Without them, financial markets would not be able to move funds from people who save to people who have productive investment opportunities. Thus financial institutions play a crucial role in the economy.

Structure of the Financial System

The financial system is complex, comprising many different types of private-sector financial institutions, including banks, insurance companies, mutual funds, finance companies, and investment banks, all of which are heavily regulated by the government. If an individual wanted to make a loan to Bombardier or Blackberry, for example, they would not go directly to the president of the company and offer a loan. Instead, the individual would lend to such a company indirectly through **financial intermediaries**, which are institutions that borrow funds from people who have saved and in turn make loans to people who need funds.

Why are financial intermediaries so crucial to well-functioning financial markets? Why do they extend credit to one party but not to another? Why do they usually write complicated legal documents when they extend loans? Why are they the most heavily regulated businesses in the economy?

We answer these questions in Chapter 8 by developing a coherent framework for analyzing financial structure in Canada and in the rest of the world.

Banks and Other Financial Institutions

Banks are financial institutions that accept deposits and make loans. The term *banks* includes firms such as chartered banks, trust and loan companies, and credit unions and *caisses populaires*. Banks are the financial intermediaries that the average person interacts with most frequently. A person who needs a loan to buy a house or a car usually obtains it from a local bank. Most Canadians keep a large portion of their financial wealth in banks in the form of chequing accounts, savings accounts, or other types of bank deposits. Because banks are the largest financial intermediaries in our economy, they deserve the most careful study. However, banks are not the only important financial institutions. Indeed, in recent years other financial institutions, such as insurance companies, finance companies, pension funds, mutual funds, and investment banks, have been growing at the expense of banks, so we need to study them as well.

In Chapter 9, we extend the economic analysis in Chapter 8 to understand why financial regulation takes the form it does and what can go wrong in the regulatory process. In Chapter 10, we look at the banking industry and examine how the competitive environment has changed this industry. We also learn why some financial institutions have been growing at the expense of others. In Chapter 12, we examine how banks and other financial institutions manage their assets and liabilities to make profits. Because the economic environment for banks and other financial institutions has become increasingly risky, these institutions must find ways to manage risk. How financial institutions manage risk with financial derivatives is the topic of Chapter 13.

Financial Innovation

In Chapter 11, we also study **financial innovation**, the development of new financial products and services. We will see why and how financial innovation takes place, with particular emphasis on how the dramatic improvements in information technology have led to new financial products and the ability to deliver financial services electronically through what has become known as **e-finance**. We also study financial innovation because it shows us how creative thinking on the part of financial institutions can

lead to higher profits but can also sometimes result in financial disasters. By studying how financial institutions have been creative in the past, we obtain a better grasp of how they may be creative in the future. This knowledge provides us with useful clues about how the financial system may change over time.

Financial Crises

At times, the financial system seizes up and produces **financial crises**, which are major disruptions in financial markets that are characterized by sharp declines in asset prices and the failures of many financial and nonfinancial firms. Financial crises have been a feature of capitalist economies for hundreds of years and are typically followed by severe business cycle downturns. Starting in August 2007, the U.S. economy was hit by the worst financial crisis since the Great Depression. Defaults in subprime residential mortgages led to major losses in financial institutions, producing not only numerous bank failures but also the demise of Bear Stearns and Lehman Brothers, two of the largest investment banks in the United States. The crisis produced the worst economic downturn since the Great Depression, and as a result, it is now referred to as the "Great Recession."

We discuss why these crises occur and why they do so much damage to the economy in Chapter 11.

1.3 WHY STUDY MONEY AND MONETARY POLICY?

LO 1.3 Identify the basic links among monetary policy, the business cycle, and economic variables.

Money, also referred to as the **money supply**, is defined as anything that is generally accepted as payment for goods or services or in the repayment of debts. Money is linked to changes in economic variables that affect all of us and are important to the health of the economy. The final two parts of this book examine the role of money in the economy.

Money and Business Cycles

During 1981–1982, the total production of goods and services (called **aggregate output**) in the Canadian economy fell and the **unemployment rate** (the percentage of the available labour force unemployed) rose to close to 12%. After 1982, the economy began to expand rapidly, and by 1989, the unemployment rate had declined to 7.5%. In 1990, the eight-year expansion came to an end, with the unemployment rate rising to above 11%. The economy bottomed out in 1991, and the subsequent recovery was the longest in Canadian history up to that time, with the unemployment rate falling to around 6% in 2008 before rising to over 8% in early 2009 in the aftermath of the global financial crisis. By early 2020, the unemployment rate had fallen to 3.5%, only to rise sharply starting in March 2020 with the onset of the Covid-19 recession.

Why did the economy undergo such pronounced fluctuations? Evidence suggests that money plays an important role in generating **business cycles**, the upward and downward movement of aggregate output produced in the economy. Business cycles affect all of us in immediate and important ways. When output is rising, for example, it is easier to find a good job; when output is falling, finding a good job might be difficult. Figure 1-3 shows the movements of the rate of growth of the money supply over

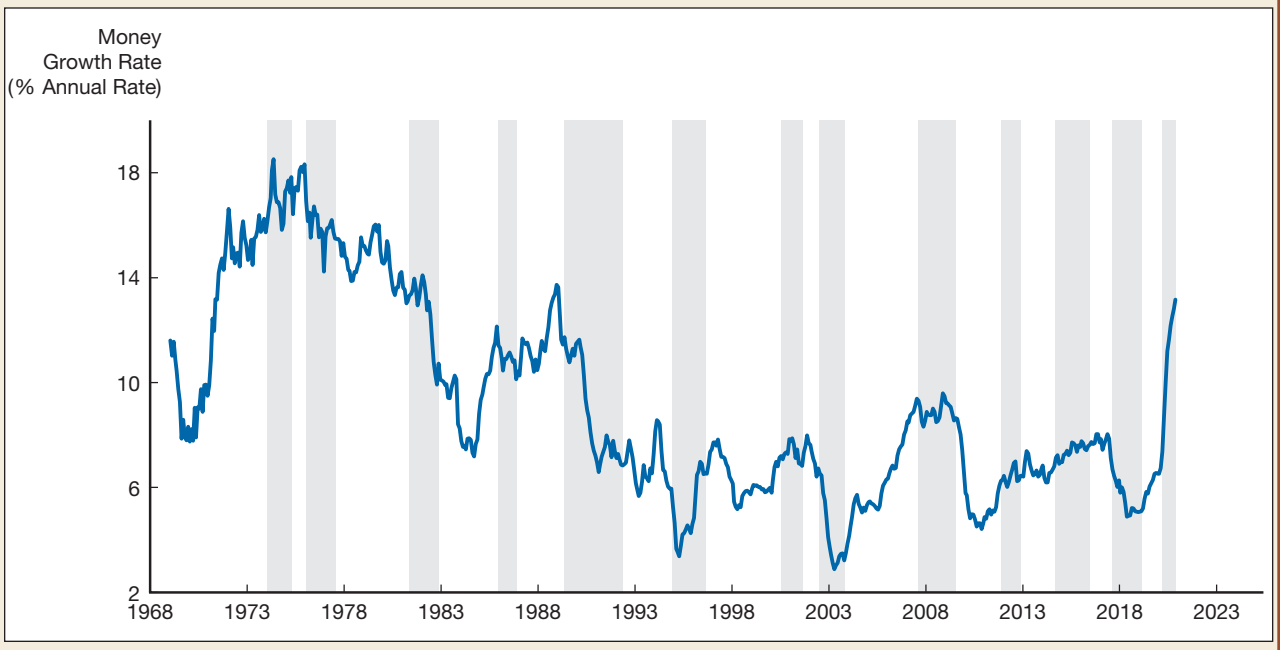


FIGURE 1-3 Money Growth (M1++ [gross] Annual Rate) and the Business Cycle in Canada, 1968–2020
Although money growth has declined before almost every recession, not every decline in the rate of money growth is followed by a recession. Shaded areas represent recessions.
Source: Statistics Canada CANSIM series v41552801.

the 1968–2020 period, with the shaded areas representing **recessions**, or periods of declining aggregate output. We see that the rate of money growth declined before most recessions, indicating that changes in money growth might be a driving force behind business cycle fluctuations. However, declines in the rate of money growth are often not followed by a recession.

We explore how money and monetary policy might affect aggregate output in Chapters 20 through 26 (Part 7) of this book, where we study **monetary theory**, the theory that relates the quantity of money and monetary policy to changes in aggregate economic activity and inflation.

Money and Inflation

The movie you paid $10 to see last week would have set you back only a dollar or two 30 years ago. In fact, for $10, you probably could have had dinner, seen the movie, and bought yourself a big bucket of hot buttered popcorn. As shown in Figure 1-4, which illustrates the movement of average prices in the Canadian economy from 1970 to 2020, the prices of most items are quite a bit higher now than they were then. The average price of goods and services in an economy is called the **aggregate price level** or, more simply,

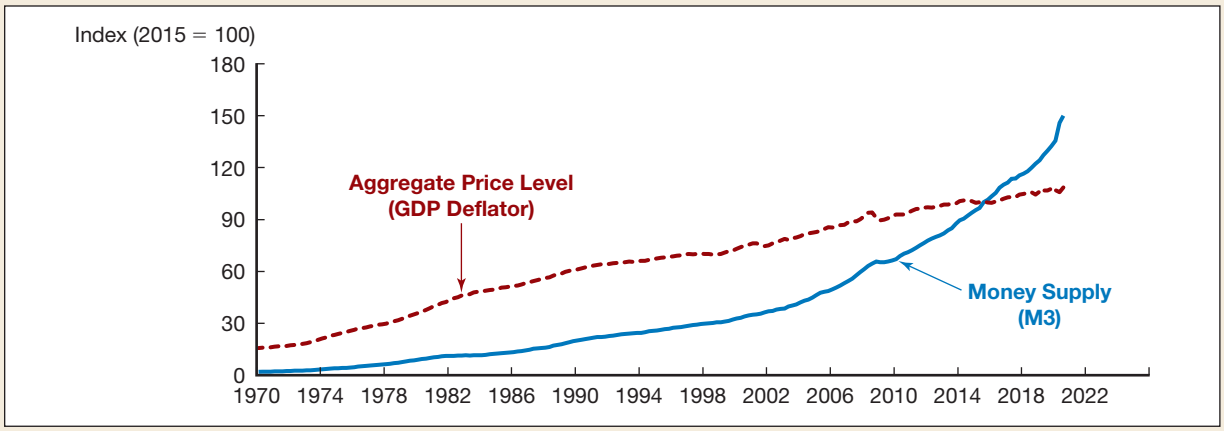


FIGURE 1-4 Aggregate Price Level and the Money Supply in Canada, 1970–2020
From 1970 to 2020, the price level has increased more than sevenfold.
Source: Federal Reserve Bank of St. Louis, FRED database: https://fred.stlouisfed.org/series/M2SL; https://fred.stlouisfed.org/series/GDPDEF

the *price level* (a more precise definition is found in the appendix to this chapter). From 1970 to 2020, the price level has increased more than sevenfold. **Inflation**, a continual increase in the price level, affects individuals, businesses, and the government. It is generally regarded as an important problem to be solved and is often at the top of political and policymaking agendas. To solve the inflation problem, we need to know something about its causes.

What explains inflation? One clue to answering this question is found in Figure 1-4, which plots the money supply versus the price level. As we can see, the price level and the money supply generally rise together. These data seem to indicate that a continuing increase in the money supply might be an important factor in causing the continuing increase in the price level that we call inflation.

Further evidence that inflation may be tied to continuing increases in the money supply is found in Figure 1-5, which plots the average **inflation rate** (the rate of change of the price level, usually measured as a percentage change per year) for a number of countries over the 10-year period 2009–2019 against the average rate of money growth over the same period. As you can see, a positive association exists between inflation and the growth rate of the money supply: The countries with high money growth rates, such as Russia and Turkey, tend to have higher inflation rates. By contrast, Japan and the Euro area experienced low inflation rates over the same period, and their rates of money growth were low. Such evidence led Milton Friedman, a Nobel laureate in economics, to make the famous statement, "Inflation is always and everywhere a monetary phenomenon."[2] We look at the quantity of money and monetary policy's role in creating inflation in Chapters 20 and 24.

[2]Milton Friedman, *Dollars and Deficits* (Upper Saddle River, NJ: Prentice Hall, 1968), 39.

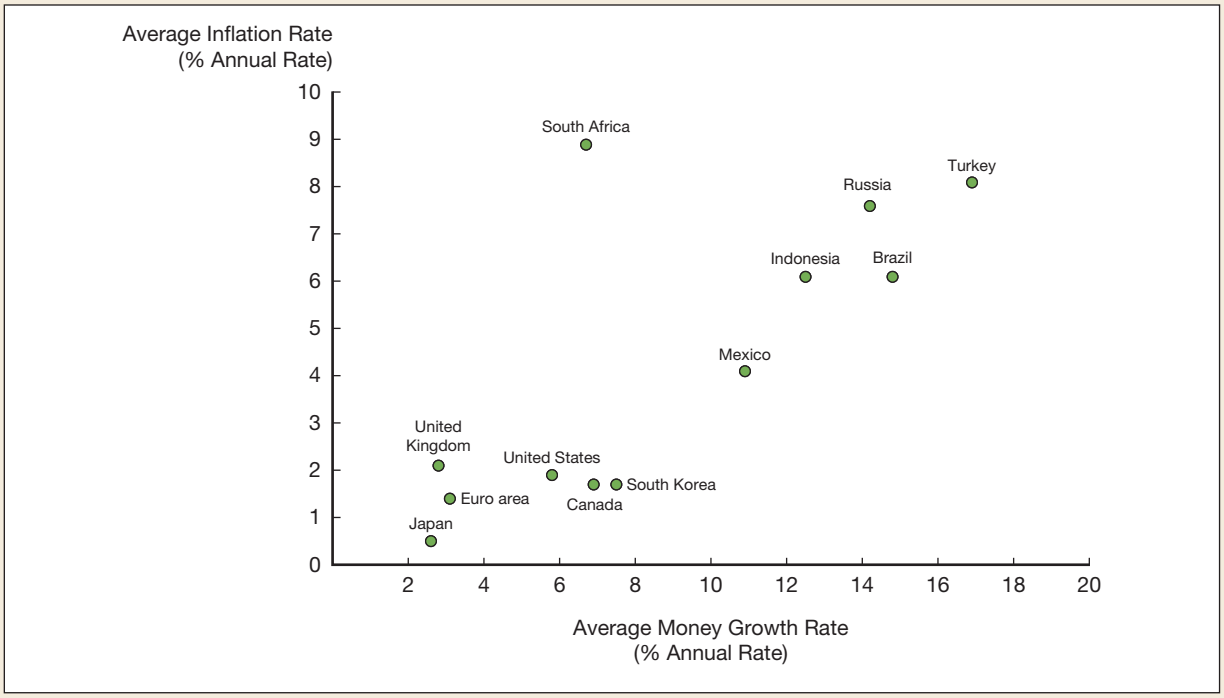


FIGURE 1-5 Average Inflation Rate Versus Average Rate of Money Growth for Selected Countries, 2009–2019

A positive association can be seen between the 10-year averages of inflation and the growth rate of the money supply: Countries with high money growth rates tend to have higher inflation rates.

Source: Federal Reserve Bank of St. Louis, FRED database: fred.stlouisfed.org/

Money and Interest Rates

In addition to other factors, money plays an important role in interest-rate fluctuations, which are of great concern to businesses and consumers. Figure 1-6 shows changes in the interest rate on long-term Canada bonds and the rate of money growth from 1976 to 2020. As the money growth rate rose in the late 1970s, the long-term bond rate rose with it. However, the relationship between money growth and interest rates has been less clear-cut since 1980. We analyze the relationship between money growth and interest rates when we examine the behaviour of interest rates in Chapter 5.

Conduct of Monetary Policy

Because money affects many economic variables that are important to the well-being of our economy, politicians and policymakers throughout the world care about the conduct of **monetary policy**, the management of money and interest rates. The organization responsible for the conduct of a nation's monetary policy is the **central bank**. Canada's central bank is the **Bank of Canada** (also called simply "**the Bank**"). In Chapters 14 through 17 (Part 5), we study how central banks such as the Bank of

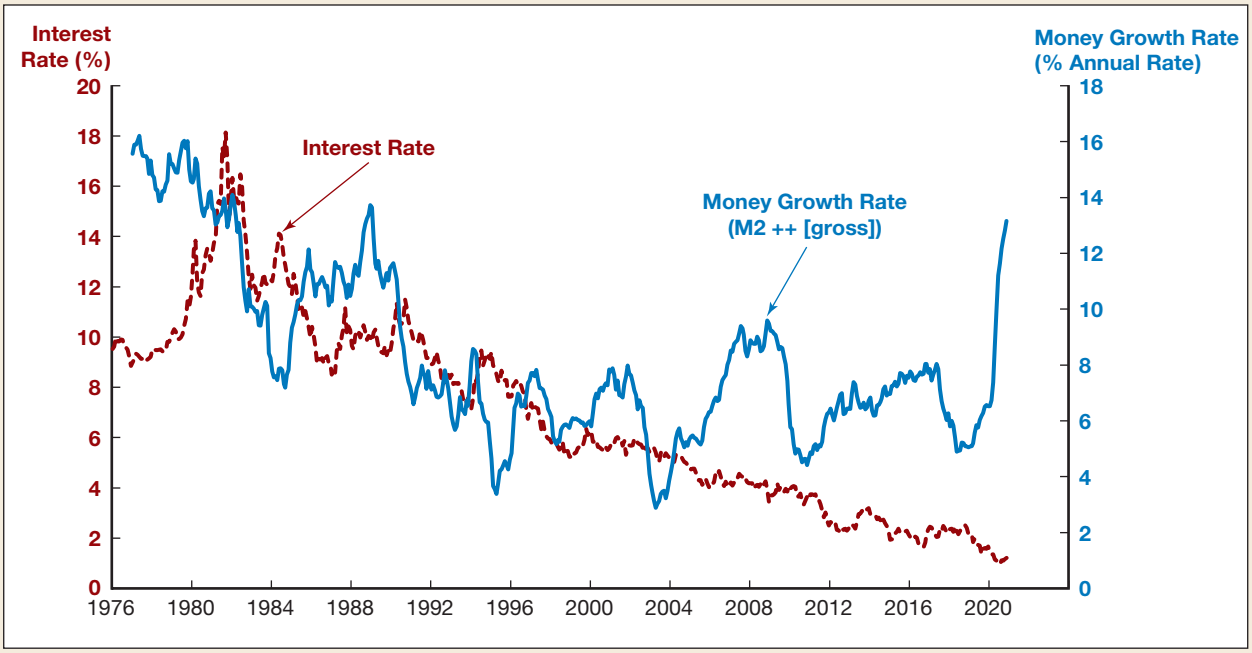


FIGURE 1-6 **Money Growth (M2++ [gross] Annual Rate) and Interest Rates (Long-Term Government of Canada Bonds), 1976–2020**

As the money growth rate rose in the 1980s and 1990s, the long-term bond rate rose with it. However, the relationship between money growth and interest rates has been less clear-cut since 2000.

Source: Statistics Canada CANSIM series V41552801 and V122544.

Canada can affect the quantity of money and interest rates in the economy, and then we look at how monetary policy is actually conducted in Canada and elsewhere.

Fiscal Policy and Monetary Policy

Fiscal policy involves decisions about government spending and taxation. A **budget deficit** is an excess of government expenditures with respect to tax revenues for a particular time period, typically a year, while a **budget surplus** arises when tax revenues exceed government expenditures. The government must finance any budget deficit by borrowing, whereas a budget surplus leads to a lower government debt burden. As Figure 1-7 shows, the budget deficit, relative to the size of the Canadian economy, peaked in 1982 at 8.7% of national output (as calculated by the **gross domestic product**, or **GDP**, a measure of aggregate output described in the appendix to this chapter). Since then, the budget deficit at first declined, and the budget actually went into surplus from 1997 to 2008. In the aftermath of the 2007–2009 global financial crisis, the budget has swung back again into deficit, with recent deficits close to 4% of GDP and then falling substantially thereafter. The budget deficit then rose starting in 2016 and shot up sharply when the coronavirus pandemic hit the economy in 2020. What to do about budget deficits has been the subject of bitter battles between Canadian politicians.

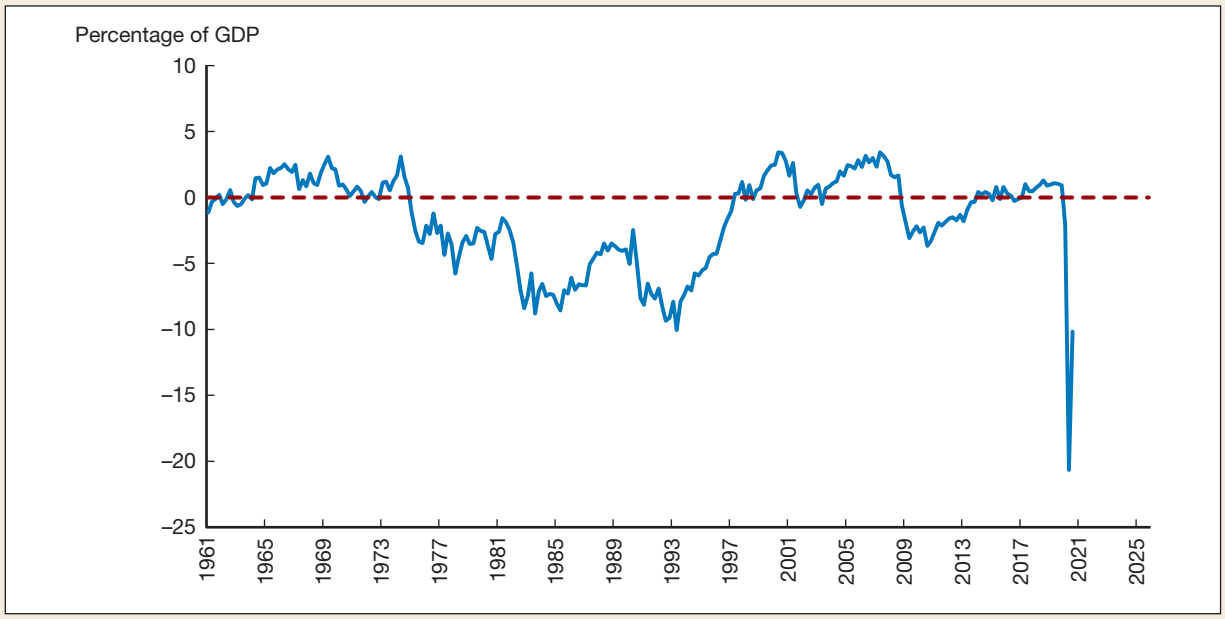


FIGURE 1-7 Government Budget Surplus or Deficit as a Percentage of Gross Domestic Product, 1961–2020

The budget deficit, relative to the size of the Canadian economy, has fluctuated substantially over the years. It rose to over 8% of GDP in 1993 and then fell, eventually leading to budget surpluses from 1997 to 2008. Subsequently, budget deficits climbed, peaking at over 3% of GDP in 2009, fell substantially thereafter, and then rose starting in 2016.

Source: Statistics Canada CANSIM series V62425528, V62425550, and V62305783.

You may have read statements in newspapers or heard on TV that budget surpluses are a good thing, while deficits are undesirable. In Chapter 20, we examine why deficits might result in a higher rate of money growth, a higher rate of inflation, and higher interest rates.

1.4 WHY STUDY INTERNATIONAL FINANCE?

LO 1.4 Explain the importance of exchange rates in a global economy.

The globalization of financial markets has accelerated at a rapid pace in recent years. Financial markets have become increasingly integrated throughout the world. Canadian companies often borrow in foreign financial markets, and foreign companies borrow in Canadian financial markets. Banks and other financial institutions have become increasingly international, with operations in many countries throughout the world. Part 6 of this book explores the foreign exchange market and the international financial system.

The Foreign Exchange Market

For funds to be transferred from one country to another, they have to be converted from the currency of the country of origin (say, dollars) into the currency of the country they are going to (say, euros). The **foreign exchange market** is where this conversion takes place, so it is instrumental in moving funds between countries. It is also important because it is where the **foreign exchange rate**, or the price of one country's currency in terms of another's, is determined.

Figure 1-8 shows the exchange rate for the Canadian dollar from 1971 to 2020 in terms of the U.S. dollar. The exchange rate is defined as the U.S. dollar price of one Canadian dollar. Clearly, the exchange rate has experienced five long swings over this period. The first is a 30.8% depreciation from January 1973 to February 1986; the second is a 26% appreciation from February 1986 to January 1992; the third is a 26.7% depreciation from January 1992 to October 2002; the fourth is a 54% appreciation from October 2002 to May 2008; and the fifth is a depreciation of close to 32% since then.

What have these fluctuations in the exchange rate meant to the Canadian public and businesses? A change in the exchange rate has a direct effect on Canadian consumers because it affects the cost of imports. In particular, a weaker dollar leads to more expensive foreign goods, makes vacationing abroad more expensive, and raises the cost of indulging your desire for imported delicacies. When the value of the dollar drops, Canadians decrease their purchases of foreign goods and increase their consumption of domestic goods (such as travel within Canada or Canadian-made wine).

FIGURE 1-8 **Exchange Rate of the Canadian Dollar, 1971–2020**

The value of the Canadian dollar relative to the U.S. dollar has fluctuated substantially over the years.

Source: Federal Reserve Bank of St. Louis, FRED database: series EXCAUS

Conversely, a strong dollar means that Canadian goods exported abroad will cost more in foreign countries, and hence foreigners will buy fewer of them. Exports of steel, for example, declined sharply when the dollar strengthened in the late 1980s. A strong dollar benefited Canadian consumers by making foreign goods cheaper but hurt Canadian businesses and eliminated some jobs by cutting both domestic and foreign sales of the businesses' products. The increase in the value of the dollar since 2002 has had a similar effect: It made foreign goods less expensive but made Canadian businesses less competitive. Fluctuations in the foreign exchange markets have major consequences for the Canadian economy.

In Chapter 18, we study how exchange rates are determined in the foreign exchange market, in which Canadian dollars are bought and sold for foreign currencies.

The International Financial System

The tremendous increase in capital flows among countries has heightened the international financial system's impact on domestic economies. Issues we will explore in Chapter 19 include

- How does a country's decision to fix its exchange rate to that of another nation shape the conduct of monetary policy?
- What is the impact of capital controls that restrict mobility of capital across national borders on domestic financial systems and the performance of the economy?
- What role should international financial institutions, such as the International Monetary Fund, play in the international financial system?

1.5 MONEY, BANKING, AND FINANCIAL MARKETS AND YOUR CAREER

LO 1.5 Explain how the study of money, banking, and financial markets may advance your career.

Before taking this class, you might have asked yourself the practical question, "How will the study of money, banking, and financial markets help my career?" For some of you, the answer is straightforward. Financial institutions are among the largest employers in the country, and studying money, banking, and financial markets can help you get a good job in the financial sector.

Even if your interests lie elsewhere, the study of money, banking, and financial institutions can help advance your career because at many times in your life, as an employee or the owner of a business, the critical thinking skills learned in this study will improve your performance. For example, understanding monetary policy may help you predict when interest rates will rise or fall, helping you to make decisions about whether it is better to borrow now or to wait until later. Knowing how banks and other financial institutions are managed may help you get a better deal when you need to borrow from them or if you decide to supply them with funds. Knowledge of how financial markets work may enable you to make better investment decisions, whether for yourself or for the company you work for.

1.6 HOW WE WILL STUDY MONEY, BANKING, AND FINANCIAL MARKETS

LO 1.6 Describe how the text approaches the teaching of money, banking, and financial markets.

This textbook stresses the "economic way of thinking" by developing a unifying framework in which you will study money, banking, and financial markets. This analytic framework uses a few basic economic concepts to organize your thinking about the determination of asset prices, the structure of financial markets, bank management, and the role of money in the economy. It encompasses the following basic concepts:

- A simplified approach to the demand for assets
- The concept of equilibrium
- Basic supply and demand analysis to explain behaviour of financial markets
- The search for profits
- An approach to financial structure based on transaction costs and asymmetric information
- Aggregate supply and demand analysis

The unifying framework used in this book will keep your knowledge from becoming obsolete and make the material more interesting. It will enable you to learn what *really* matters without having to memorize a mass of dull facts that you will forget soon after the final exam. This framework will also provide you with the tools you need to understand trends in the financial marketplace and in variables such as interest rates, exchange rates, inflation, and aggregate output.

To help you understand and apply the unifying analytic framework, simple models are constructed in which the variables held constant are carefully delineated. Each step in the derivation of the model is clearly and carefully laid out, and the models are then used to explain various phenomena by focusing on changes in one variable at a time, holding all other variables constant.

To reinforce the models' usefulness, this text uses case studies, applications, and special-interest boxes to present evidence that supports or casts doubt on the theories being discussed. This exposure to real-life events and empirical data should dissuade you from thinking that all economists do is make abstract assumptions and develop theories that have little to do with actual behaviour.

To function better financially in the real world outside the classroom, you must have the tools with which to follow the financial news that is reported in leading financial publications and on the web. To help and encourage you to read the financial news, this book contains special boxed inserts titled "Following the Financial News" that provide detailed information and definitions to help you evaluate data that are discussed frequently in the media. This text also allows you to view the most current data for a high percentage of the in-text data figures using the Statistics Canada CANSIM database and the Federal Reserve Bank of St. Louis's FRED database. Figures for which you can do this are labelled *Real-Time Data.*

To master any field, you need to practise, practise, practise. To help you in this endeavour, this book contains over 700 end-of-chapter questions and applied problems that ask you to apply the analytic concepts you have learned to real-world issues. In addition, at the end of almost every chapter there are several real-time data analysis problems, which ask you to download the most recent data from the Statistics Canada CANSIM database and the Federal Reserve Bank of St. Louis's FRED database and then use these data to answer interesting questions.

CONCLUDING REMARKS

The study of money, banking, and financial markets is an exciting field that directly affects your life and career. Interest rates influence the earnings you make on your savings and the payments on loans you may seek for a car or a house, and monetary policy may affect your job prospects and the prices you will pay for goods in the future. Your study of money, banking, and financial markets will introduce you to many of the controversies related to economic policy that are hotly debated in the political arena and will help you gain a clearer understanding of the economic phenomena you hear about in the news media. The knowledge you gain will stay with you and benefit you long after this course is over.

SUMMARY

1. Activities in financial markets directly affect individuals' wealth, the behaviour of businesses, and the efficiency of our economy. Three financial markets deserve particular attention: the bond market (where interest rates are determined), the stock market (which has a major effect on people's wealth and on firms' investment decisions), and the foreign exchange market (because fluctuations in the foreign exchange rate have major consequences for the Canadian economy).
2. Banks and other financial institutions channel funds from people who might not put them to productive use to people who can do so and thus play a crucial role in improving the efficiency of the economy. When the financial system seizes up and produces a financial crisis, financial firms fail, which causes severe damage to the economy.
3. Money and monetary policy appear to have a major influence on inflation, business cycles, and interest rates. Because these economic variables are so important to the health of the economy, we need to understand how monetary policy is and should be conducted. We also need to study government fiscal policy because it can be an influential factor in the conduct of monetary policy.
4. The study of money, banking, and financial markets can help advance your career by helping you get a high-paying job in the financial sector, decide when you or your firm should borrow, get a better deal from financial institutions, or make better investment decisions.
5. This textbook stresses the "economic way of thinking" by developing a unifying analytic framework in which to study money, banking, and financial markets, using a few basic economic principles. The textbook also emphasizes the interaction of theoretical analysis and empirical data.

MyLab Economics Make the grade with MyLab Economics: Select end-of-chapter questions can be found on MyLab Economics. You can practise them as often as you want, and most feature step-by-step guided instructions to help you find the right answer.

QUESTIONS

1. What is the typical relationship among interest rates on three-month Treasury bills, long-term Government of Canada bonds, and corporate bonds?
2. What effect might a fall in stock prices have on business investment?
3. Explain the main difference between a bond and a common stock.
4. Explain the link between well-performing financial markets and economic growth. Name one channel through which financial markets might affect economic growth and poverty.
5. What was the main cause of the global financial crisis that began in 2007?

6. Can you think of a reason why people in general do not lend money to one another to buy a house or a car? How would your answer explain the existence of banks?
7. What are the other important financial intermediaries in the economy besides banks?
8. Can you date the latest financial crisis in the United States or in Europe? Are there reasons to think that these crises might have been related? Why?
9. Has the inflation rate in Canada increased or decreased in the past few years? What about interest rates?
10. If history repeats itself and we see a decline in the rate of money growth, what might you expect to happen to
    a. real output?
    b. the inflation rate?
    c. interest rates?
11. When interest rates decrease, how might businesses and consumers change their economic behaviour?
12. Is everybody worse off when interest rates rise?
13. Why do managers of financial institutions care so much about the activities of the Bank of Canada?
14. How does the current size of the Canadian budget deficit compare to the historical budget deficit or surplus for the time period since 1960?
15. How would a fall in the value of the pound sterling affect British consumers?
16. How would an increase in the value of the pound sterling affect Canadian businesses?
17. How can changes in foreign exchange rates affect the profitability of financial institutions?
18. According to Figure 1-8, in which years would you have chosen to visit the Canadian Rockies rather than the Grand Canyon in Arizona?
19. When the dollar is worth more in relation to currencies of other countries, are you more likely to buy Canadian-made or foreign-made jeans? Are Canadian companies that manufacture jeans happier when the dollar is strong or when it is weak? What about a Canadian company that is in the business of importing jeans into Canada?
20. Much of the Canadian government debt is held by foreign investors as Government of Canada bonds and bills. How do fluctuations in the dollar exchange rate affect the value of that debt held by foreigners?

APPLIED PROBLEMS

21. The following table lists the foreign exchange rate between U.S. dollars and British pounds (GBP) during May 2020. Which day would have been the best for converting US$200 into British pounds? Which day would have been the worst? What would be the difference in pounds?

Date	$/£	Date	$/£
05-01	1.2509	05-18	1.2211
05-04	1.2430	05-19	1.2255
05-05	1.2449	05-20	1.2257
05-06	1.2347	05-21	1.2227
05-07	1.2349	05-22	1.2178
05-08	1.2436	05-25	NA
05-11	1.2330	05-26	1.2337
05-12	1.2299	05-27	1.2231
05-13	1.2225	05-28	1.2325
05-14	1.2194	05-29	1.2320
05-15	1.2129		

DATA ANALYSIS PROBLEMS

The Problems update with real-time data in **MyLab Economics** and are available for practice or instructor assignment.

1. Go to the Statistics Canada CANSIM database and find data on the M2++ (gross) money supply and the 10-year Canada bond rate from January 2000 to November 2020. Add the two series into a single graph. Transform the M2++ money supply variable into the M2++ growth rate by adjusting the *units* for the M2++ money supply to "Percent Change from Year Ago."

a. In general, how have the growth rate of the M2++ money supply and the 10-year Canada bond rate behaved during recessions and during expansionary periods since the year 2000?

b. In general, is there an obvious, stable relationship between money growth and the 10-year interest rate since the year 2000?

c. Compare the money growth rate and the 10-year interest rate for the most recent month available to the rates for January 2000. How do the rates compare?

2. **Real-Time Data Analysis** Go to the St. Louis Federal Reserve FRED database and find data on the three-month Treasury bill rate (TB3MS), the three-month AA nonfinancial commercial paper rate (CPN3M), the 30-year Treasury bond rate (GS30), the 30-year conventional mortgage rate (MORTGAGE30US), and the NBER recession indicators (USREC). For the mortgage rate indicator, set the frequency to "monthly."

a. In general, how do these interest rates behave during recessions and during expansionary periods?

b. In general, how do the three-month rates compare to the 30-year rates? How do the Treasury rates compare to the respective commercial paper and mortgage rates?

c. For the most recent available month of data, take the average of each of the three-month rates and compare it to the average of the three-month rates from January 2000. How do the averages compare?

d. For the most recent available month of data, take the average of each of the 30-year rates and compare it to the average of the 30-year rates from January 2000. How do the averages compare?

APPENDIX TO CHAPTER 1

Defining Aggregate Output, Income, the Price Level, and the Inflation Rate

Because these terms are used so frequently throughout the text, we need to have a clear understanding of the definitions of *aggregate output*, *income*, the *price level*, and the *inflation rate*.

AGGREGATE OUTPUT AND INCOME

The most commonly reported measure of aggregate output, the **gross domestic product (GDP)**, is the market value of all final goods and services produced in a country during the course of a year. This measure excludes two sets of items that at first glance you might think it would include. Purchases of goods that have been produced in the past, whether a Rembrandt painting or a house built 20 years ago, are not counted as part of GDP; nor are purchases of stocks or bonds. Neither of these categories enters into the GDP because these categories do not include goods and services produced during the course of the year. Intermediate goods, which are used up in producing final goods and services, such as the sugar in a candy bar or the energy used to produce steel, are also not counted separately as part of the GDP. Because the value of the final goods already includes the value of the intermediate goods, to count them separately would be to count them twice.

Aggregate income, the total income of *factors of production* (land, labour, and capital) from producing goods and services in the economy during the course of the year, is best thought of as being equal to aggregate output. Because the payments for final goods and services must eventually flow back to the owners of the factors of production as income, income payments must equal payments for final goods and services. For example, if the economy has an aggregate output of $10 trillion, total income payments in the economy (aggregate income) are also $10 trillion.

REAL VERSUS NOMINAL MAGNITUDES

When the total value of final goods and services is calculated using current prices, the resulting GDP measure is referred to as *nominal GDP*. The word *nominal* indicates that values are measured using current prices. If all prices doubled but actual production of goods and services remained the same, nominal GDP would double, even though people would not enjoy the benefits of twice as many goods and services. As a result, nominal variables can be misleading measures of economic well-being.

A more reliable measure of economic production expresses values in terms of prices for an arbitrary base year, currently 2009. GDP measured with constant prices is referred to as *real GDP*, the word *real* indicating that values are measured in terms of fixed prices. Real variables thus measure the quantities of goods and services and do not change because prices have changed but rather only if actual quantities have changed.

A brief example will make the distinction clearer. Suppose that you have a nominal income of $30 000 in 2022 and that your nominal income was $15 000 in 2012. If all prices doubled between 2012 and 2022, are you better off? The answer is no: Although your income has doubled, your $30 000 buys you only the same amount of goods because prices have also doubled. A *real* income measure indicates that your income in terms of the goods it can buy is the same. Measured in 2012 prices, the $30 000 of nominal income in 2022 turns out to be only $15 000 of real income. Because your real income is actually the same for the two years, you are no better or worse off in 2022 than you were in 2012.

Because real variables measure quantities in terms of real goods and services, they are typically of more interest than nominal variables. In this text, discussion of aggregate output or aggregate income always refers to real measures (such as real GDP).

AGGREGATE PRICE LEVEL

In this chapter, we defined the aggregate price level as a measure of average prices in the economy. Three measures of the aggregate price level are commonly encountered in economic data. The first is the *GDP deflator*, which is defined as nominal GDP divided by real GDP. Thus, if 2022 nominal GDP is $10 trillion but 2022 real GDP in 2012 prices is $9 trillion,

$$\text{GDP deflator} = \frac{\$10 \text{ trillion}}{\$9 \text{ trillion}} = 1.11$$

The GDP deflator equation indicates that, on average, prices have risen 11% since 2012. Typically, measures of the price level are presented in the form of a price index, which expresses the price level for the base year (in our example, 2012) as 100. Thus the GDP deflator for 2022 would be 111.

Another popular measure of the aggregate price level (which officials in the Bank of Canada frequently focus on) is the *PCE deflator*, which is similar to the GDP deflator and is defined as nominal personal consumption expenditures (PCE) divided by real PCE.

The measure of the aggregate price level that is most frequently reported in the press is the *consumer price index* (*CPI*). The CPI is measured by pricing a "basket" of goods and services bought by a typical urban household. If, over the course of the year, the cost of this basket of goods and services rises from $500 to $600, the CPI has risen by 20%. The CPI is also expressed as a price index with the base year equal to 100.

The CPI, the PCE deflator, and the GDP deflator measures of the price level can be used to convert or deflate a nominal magnitude into a real magnitude. This is accomplished by dividing the nominal magnitude by the price index. In our example, in which the GDP deflator for 2022 is 1.11 (expressed as an index value of 111), real GDP for 2022 equals

$$\frac{\$10 \text{ trillion}}{1.11} = \$9 \text{ trillion in 2012 prices},$$

which corresponds to the real GDP figure for 2022 assumed earlier.

GROWTH RATES AND THE INFLATION RATE

The media often talk about the economy's growth rate, and particularly the growth rate of real GDP. A growth rate is defined as the percentage change in a variable, that is,

$$\text{growth rate of } x = \frac{x_t - x_{t-1}}{x_{t-1}} \times 100$$

where t indicates today and $t - 1$ indicates a year earlier.

For example, if real GDP grew from \$9 trillion in 2022 to \$9.5 trillion in 2023, then the GDP growth rate for 2023 would be 5.6%:

$$\text{GDP growth rate} = \frac{\$9.5 \text{ trillion} - \$9 \text{ trillion}}{\$9 \text{ trillion}} \times 100 = 5.6\%$$

The inflation rate is defined as the growth rate of the aggregate price level. Thus, if the GDP deflator rose from 111 in 2022 to 113 in 2023, the inflation rate using the GDP deflator would be 1.8%:

$$\text{inflation rate} = \frac{113 - 111}{111} \times 100 = 1.8\%$$

If the growth rate is for a period of less than one year, it is usually reported on an annualized basis; that is, it is converted to the growth rate over a year's time, assuming that the growth rate remains constant. For GDP, which is reported quarterly, the annualized growth rate would be approximately four times the percentage change in GDP from the previous quarter. For example, if GDP rose $\frac{1}{2}$% from the first quarter of 2022 to the second quarter of 2022, then the annualized GDP growth rate for the second quarter of 2022 would be reported as 2%($= 4 \times \frac{1}{2}$%). (A more accurate calculation would be 2.02%, because a precise quarterly growth rate should be compounded on a quarterly basis.)

2 An Overview of the Financial System

Learning Objectives

2.1 Compare and contrast direct and indirect finance.

2.2 Identify the structure and components of financial markets.

2.3 List and describe the different types of financial market instruments.

2.4 Recognize the international dimensions of financial markets.

2.5 Summarize the roles of transaction costs, risk sharing, and information costs as they relate to financial intermediaries.

2.6 List and describe the different types of financial intermediaries.

2.7 Identify the reasons for and list the types of financial market regulations.

Preview

Inez the Inventor has designed a low-cost robot that cleans the house (even does the windows!), washes the car, and mows the lawn, but she has no funds to put her wonderful invention into production. Walter the Widower has plenty of savings, which he and his wife accumulated over the years. If Inez and Walter could get together so that Walter could provide funds to Inez, Inez's robot would see the light of day, and the economy would be better off: We would have cleaner houses, shinier cars, and more beautiful lawns.

Financial markets (bond and stock markets) and financial intermediaries (such as banks, insurance companies, and pension funds) serve the basic function of getting people like Inez and Walter together so that funds can move from those who have a surplus of funds (Walter) to those who have a shortage of funds (Inez). More realistically, when Apple invents a better iPad, it may need funds to bring its new product to market. Similarly, when a local government needs to build a road or a school, it may require more funds than local property taxes provide. Well-functioning financial markets and financial intermediaries are crucial to economic health.

To study the effects of financial markets and financial intermediaries on the economy, we need to acquire an understanding of their general structure and operation. In this chapter, we learn about the major financial intermediaries and the instruments that are traded in financial markets as well as how these markets are regulated.

This chapter presents an overview of the fascinating study of financial markets and institutions. We return to a more detailed treatment of the regulation, structure, and evolution of the financial system in Chapters 8 through 11.

2.1 FUNCTION OF FINANCIAL MARKETS

LO 2.1 Compare and contrast direct and indirect finance.

Financial markets perform the essential economic function of channelling funds from households, firms, and governments that have saved surplus funds by spending less than their income to those that have a shortage of funds because they wish to spend more than their income. This function is shown schematically in Figure 2-1. Those who have saved and are lending funds, the lender-savers, are at the left, and those who

Mini-lecture

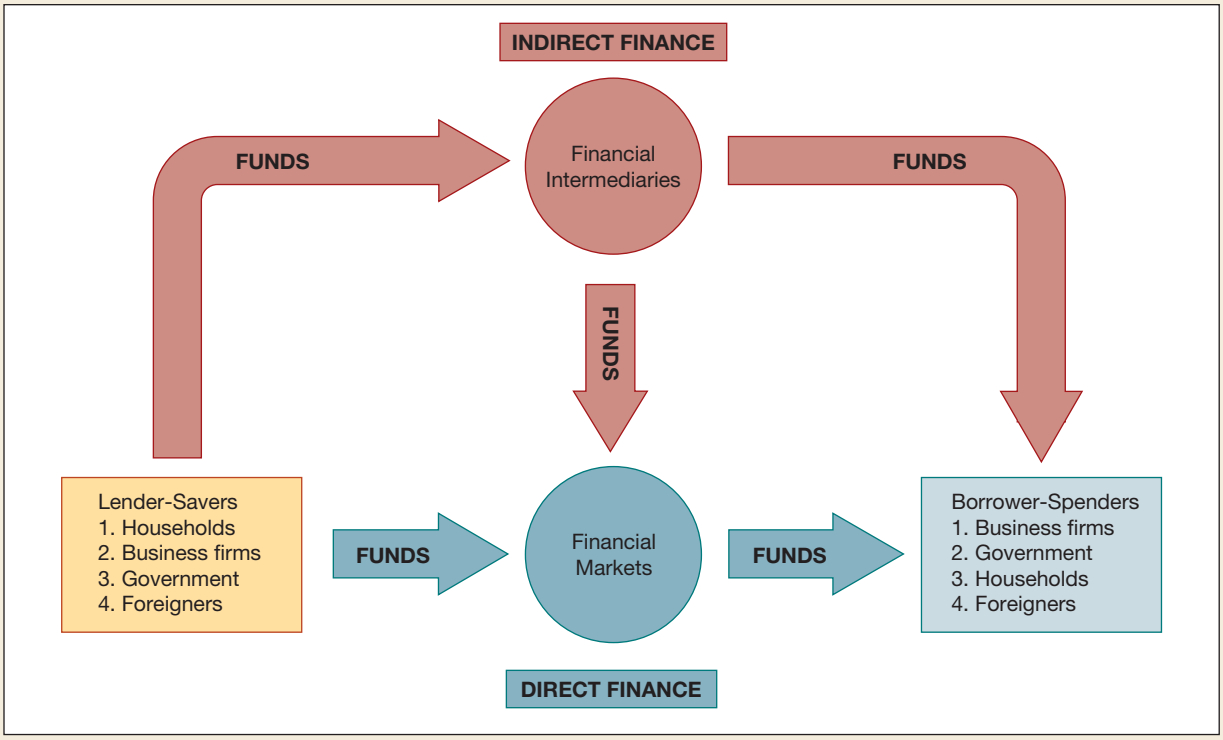


FIGURE 2-1 Flows of Funds Through the Financial System

The arrows show that funds flow from lender-savers to borrower-spenders via two routes: *direct finance,* in which borrowers borrow funds directly from financial markets by selling securities, and *indirect finance,* in which a financial intermediary borrows funds from lender-savers and then uses these funds to make loans to borrower-spenders.

must borrow funds to finance their spending, the borrower-spenders, are at the right. The principal lender-savers are households, but business enterprises and the government (particularly state and local government), as well as foreigners and their governments, sometimes also find themselves with excess funds and so lend them out. The most important borrower-spenders are businesses and the government (particularly the federal government), but households and foreigners also borrow to finance their purchases of cars, furniture, and houses. The arrows show that funds flow from lender-savers to borrower-spenders via two routes.

In *direct finance* (the route at the bottom of Figure 2-1), borrowers borrow funds directly from lenders in financial markets by selling the lenders *securities* (also called *financial instruments*), which are claims on the borrower's future income or assets. Securities are assets for the person who buys them but **liabilities** (IOUs or debts) for the individual or firm that sells (issues) them. For example, if Ford needs to borrow funds to pay for a new factory to manufacture electric cars, it might borrow the funds from savers by selling them a *bond*, a debt security that promises to make periodic payments for a specified period of time, or a *stock*, a security that entitles the owner to a share of the company's profits and assets.

Why is this channelling of funds from savers to spenders so important to the economy? The answer is that the people who save are frequently not the same people who have profitable investment opportunities available to them, the entrepreneurs. Let's first think about this on a personal level. Suppose that you have saved $1000 this year, but no borrowing or lending is possible because no financial markets are available. If you do not have an investment opportunity that will permit you to earn income with your savings, you will just hold on to the $1000 and it will earn no interest. However, Carl the Carpenter has a productive use for your $1000: He can use it to purchase a new tool that will shorten the time it takes him to build a house, thereby earning an extra $200 per year. If you could get in touch with Carl, you could lend him the $1000 at a rental fee (interest) of $100 per year, and both of you would be better off. You would earn $100 per year on your $1000, instead of the zero amount that you would earn otherwise, while Carl would earn $100 more income per year (the $200 extra earnings per year minus the $100 rental fee for the use of the funds).

In the absence of financial markets, you and Carl the Carpenter might never get together. You would both be stuck with the status quo, and both of you would be worse off. Without financial markets, it is hard to transfer funds from a person who has no investment opportunities to one who has them. Financial markets are thus essential to promoting economic efficiency.

The existence of financial markets is beneficial even if someone borrows for a purpose other than increasing production in a business. Say that you are recently married, have a good job, and want to buy a house. You earn a good salary, but because you have just started to work, you have not saved much. Over time, you would have no problem saving enough to buy the house of your dreams, but by then you would be too old to get full enjoyment from it. Without financial markets, you are stuck; you cannot buy the house and must continue to live in your tiny apartment.

If a financial market were set up so that people who had built up savings could lend you the funds to buy the house, you would be more than happy to pay them some interest so that you could own a home while you are still young enough to enjoy it. Then, over time, you would pay back your loan. If this loan could occur, you would be better off, as would the persons who made you the loan. They would now earn some interest, whereas they would not if the financial market did not exist.

Now we can see why financial markets have such an important function in the economy. They allow funds to move from people who lack productive investment opportunities to people who have such opportunities. Financial markets are critical for producing an efficient allocation of **capital** (wealth, either financial or physical, that is employed to produce more wealth), which contributes to higher production and efficiency for the overall economy. Indeed, as we will explore in Chapter 11, when financial markets break down during financial crises, as they did during the recent global financial crisis from 2007–2009, severe economic hardship results, which can sometimes lead to dangerous political instability.

Well-functioning financial markets also directly improve the well-being of consumers by allowing them to time their purchases better. They provide funds to young people to buy what they need (and will eventually be able to afford) without forcing them to wait until they have saved up the entire purchase price. Financial markets that are operating efficiently improve the economic welfare of everyone in the society.

2.2 STRUCTURE OF FINANCIAL MARKETS

LO 2.2 Identify the structure and components of financial markets.

Now that we understand the basic function of financial markets, let's look at their structure. The following descriptions of several categories of financial markets illustrate essential features of these markets.

Debt and Equity Markets

A firm or an individual can obtain funds in a financial market in two ways. The most common method is through the issuance of a debt instrument, such as a bond or a mortgage, which is a contractual agreement by the borrower to pay the holder of the instrument fixed dollar amounts at regular intervals (interest and principal payments) until a specified date (the maturity date), when a final payment is made. The **maturity** of a debt instrument is the number of years (term) until that instrument's expiration date. A debt instrument is **short term** if its maturity term is less than a year and **long term** if its maturity term is 10 years or longer. Debt instruments with a maturity term between 1 and 10 years are said to be **intermediate term**.

The second method of raising funds is through the issuance of **equities**, such as common stock, which are claims to share in the *net income* (income after expenses and taxes) and the assets of a business. If you own one share of common stock in a company that has issued one million shares, you are entitled to 1 one-millionth of the firm's net income and 1 one-millionth of the firm's assets. Equities often make periodic payments (**dividends**) to their holders and are considered long-term securities because they have no maturity date. In addition, owning stock means that you own a portion of the firm and thus have the right to vote on issues important to the firm and to elect its directors.

The main disadvantage of owning a corporation's equities rather than its debt is that an equity holder is a *residual claimant*; that is, the corporation must pay all its debt holders before it pays its equity holders. The advantage of holding equities is that equity holders benefit directly from any increases in the corporation's profitability or asset value because equities confer ownership rights on the equity holders. Debt holders do not share in this benefit, because their dollar payments are fixed. We examine the pros and cons of debt versus equity instruments in more detail in Chapter 8, which provides an economic analysis of financial structure.

The total value of equities in the United States has fluctuated between $3 and $55 trillion since 1990, depending on the prices of shares. Although the average person is more aware of the stock market than of any other financial market, the size of the debt market is often substantially larger than the size of the equities market: At the end of 2019, the value of debt instruments in the United States was $47 trillion, while the value of equities was $55 trillion.

Primary and Secondary Markets

A **primary market** is a financial market in which new issues of a security, such as a bond or a stock, are sold to initial buyers by the corporation or government agency borrowing the funds. A **secondary market** is a financial market in which securities that have been previously issued can be resold.

The primary markets for securities are not well known to the public because the selling of securities to initial buyers often takes place behind closed doors. An important

financial institution that assists in the initial sale of securities in the primary market is the **investment bank**. The investment bank does this by **underwriting** securities: It guarantees a price for a corporation's securities and then sells them to the public.

The Toronto Stock Exchange (TSX) and the TSX Venture Exchange, in which previously issued stocks are traded, are the best-known examples of secondary markets, although the bond markets, in which previously issued bonds of major corporations and the Canadian government are bought and sold, actually have a larger trading volume. Other examples of secondary markets are foreign exchange markets, futures markets, and options markets. Securities brokers and dealers are crucial to a well-functioning secondary market. **Brokers** are agents of investors who match buyers with sellers of securities; **dealers** link buyers and sellers by buying and selling securities at stated prices.

When an individual buys a security in the secondary market, the person who has sold the security receives money in exchange for the security, but the corporation that issued the security acquires no new funds. A corporation acquires new funds only when its securities are first sold in the primary market. Nonetheless, secondary markets serve two important functions. First, they make it easier and quicker to sell these financial instruments to raise cash; that is, they make the financial instruments more **liquid**. The increased liquidity of these instruments then makes them more desirable and thus easier for the issuing firm to sell in the primary market. Second, secondary markets determine the price of the security that the issuing firm sells in the primary market. The investors who buy securities in the primary market will pay the issuing corporation no more than the price they think the secondary market will set for this security. The higher the security's price in the secondary market, the higher the price the issuing firm will receive for a new security in the primary market, and hence the greater the amount of financial capital it can raise. Conditions in the secondary market are therefore the most relevant to corporations issuing securities. For this reason, books like this one, which deal with financial markets, focus on the behaviour of secondary markets rather than that of primary markets.

Exchanges and Over-the-Counter Markets

Secondary markets can be organized in two ways. One method is through **exchanges**, where buyers and sellers of securities (or their agents or brokers) meet in one central location to conduct trades. The Toronto Stock Exchange for stocks and ICE Futures Canada (formerly Winnipeg Commodity Exchange) for commodities (wheat, oats, barley, and other agricultural commodities) are examples of organized exchanges. The Montreal Exchange (MX) is another example of an organized exchange, offering a range of equity, interest rate, and index derivative products.

The other forum for a secondary market is an **over-the-counter (OTC) market**, in which dealers at different locations who have an inventory of securities stand ready to buy and sell securities "over the counter" to anyone who comes to them and is willing to accept their prices. Because over-the-counter dealers are in contact via computers and know the prices set by one another, the OTC market is very competitive and not very different from a market with an organized exchange.

Many common stocks are traded over the counter, although a majority of the largest corporations have their shares traded at organized stock exchanges. The Canadian government bond market, by contrast, is set up as an over-the-counter market. Dealers establish a "market" in these securities by standing ready to buy and sell Canadian government bonds. Other over-the-counter markets include those that trade other types of financial instruments, such as negotiable certificates of deposit, federal funds, and foreign exchange instruments.

Money and Capital Markets

Another way of distinguishing between markets is on the basis of the maturity of the securities traded in each market. The **money market** is a financial market in which only short-term debt instruments (generally those with original maturity terms of less than one year) are traded; the **capital market** is the market in which longer-term debt instruments (generally those with original maturity terms of one year or greater) and equity instruments are traded. Money market securities are usually more widely traded than longer-term securities and so tend to be more liquid. In addition, as we will see in Chapter 4, short-term securities have smaller fluctuations in prices than long-term securities, making them safer investments. As a result, corporations and banks actively use the money market to earn interest on surplus funds that they expect to have only temporarily. Capital market securities, such as stocks and long-term bonds, are often held by financial intermediaries such as insurance companies and pension funds, which have little uncertainty about the amount of funds they will have available in the future.

2.3 FINANCIAL MARKET INSTRUMENTS

LO 2.3 List and describe the different types of financial market instruments.

To complete our understanding of how financial markets perform the important role of channelling funds from lender-savers to borrower-spenders, we need to examine the securities (instruments) traded in financial markets. We first focus on the instruments traded in the money market and then turn to those traded in the capital market.

Money Market Instruments

Because of their short terms to maturity, the debt instruments traded in the money market undergo the least price fluctuations and so are the least risky investments. The money market has undergone great changes in the past three decades, with the amounts of some financial instruments growing at a far more rapid rate than others.

The principal money market instruments are listed in Table 2-1, along with the amount at the end of 1990, 2000, 2010, and 2020.

TABLE 2-1 Principal Money Market Instruments

	Amount Outstanding ($ Millions)			
Type of Instrument	**1990**	**2000**	**2010**	**2020**
Treasury bills				
Government of Canada	113 654	76 633	150 830	190 431
Corporate short-term paper				
Commercial paper	31 303	125 751	56 338	55 766
Bankers' acceptances	46 059	51 409	47 316	88 453

Source: Statistics Canada CANSIM series V37355, V122652, and V122243.

Following the Financial News — Money Market Rates

The four money market interest rates discussed most frequently in the media are as follows:

Prime rate: The base interest rate on corporate bank loans, an indicator of the cost of businesses borrowing from banks

Overnight interest rate: The interest rate charged on overnight loans in the overnight funds market, a sensitive indicator of the cost to banks of borrowing funds from other banks and the stance of monetary policy

Treasury bill rate: The interest rate on government of Canada Treasury bills, an indicator of general interest-rate movements

Libor rate: A benchmark rate for new (short-term) lending by global banks in the international interbank market. After the end of the London Interbank Offered Rate (*LIBOR*) on December 31, 2021, the *Secured Overnight Funding Rate* (*SOFR*) will most likely be that rate.

The data for these interest rates are reported daily in newspapers and on Internet sites such as http://www.bankrate.com.

Government of Canada Treasury Bills

These short-term debt instruments of the Canadian government are issued in 1-, 3-, 6-, and 12-month maturities to finance the federal government. They pay a set amount at maturity and have no interest payments, but they effectively pay interest by initially selling at a discount—that is, at a price lower than the set amount paid at maturity. For instance, in May 2023, you might buy a six-month Treasury bill for $9000 that can be redeemed in November 2023 for $10 000.

Treasury bills are the most liquid of all money market instruments because they are the most actively traded. They are also the safest money market instrument because there is a low probability of **default**, a situation in which the party issuing the debt instrument (the federal government, in this case) is unable to make interest payments or pay off the amount owed when the instrument matures. The federal government can always meet its debt obligations because it can raise taxes or issue **currency** (paper money or coins) to pay off its debts. Treasury bills are held mainly by banks, although small amounts are held by households, corporations, and other financial intermediaries.

Certificates of Deposit

A **certificate of deposit (CD)**, also known as a fixed deposit, is a debt instrument sold by a bank to depositors that pays annual interest of a given amount and at maturity pays back the original purchase price. CDs are often negotiable, meaning that they can be traded, and in bearer form (called *bearer deposit notes*), meaning that the buyer's name is recorded neither in the issuer's books nor on the security itself. These negotiable CDs are issued in multiples of $100 000 and with maturities of 90 days to one year and can be resold in a secondary market, thus offering the purchaser both yield and liquidity.

Chartered banks also issue non-negotiable CDs. That is, they cannot be sold to someone else and cannot be redeemed from the bank before maturity without paying a substantial penalty. Non-negotiable CDs are issued in denominations ranging from $5000 to $100 000 and with maturities of one day to five years. They are also known as *term deposit receipts* or *term notes*.

CDs are also an extremely important source of funds for trust and loan companies. These institutions issue CDs under a variety of names; for example, DRs (deposit receipts), GTCs (guaranteed trust certificates), GICs (guaranteed investment certificates), and GIRs (guaranteed investment receipts).

Commercial Paper

Commercial paper is a short-term debt instrument issued in either Canadian dollars or other currencies by large banks and well-known corporations, such as General Motors and Bombardier. Because commercial paper is unsecured, only the largest and most creditworthy corporations issue commercial paper. The interest rate the corporation is charged reflects the firm's level of risk. The interest rate on commercial paper is low relative to those on other corporate fixed-income securities and slightly higher than rates on Government of Canada Treasury bills.

Sales finance companies also issue short-term promissory notes known as **finance paper**. Finance and commercial paper are issued in minimum denominations of $50 000 and in maturities of 30 to 365 days for finance paper and 1 to 365 days for commercial paper. Most finance and commercial paper is issued on a discounted basis.

Repurchase Agreements

Repurchase agreements (*repos*) are effectively short-term loans (usually with a maturity term of less than two weeks) for which Treasury bills serve as *collateral*, an asset that the lender receives if the borrower does not pay back the loan. Repos are made as follows: A large corporation, such as Bombardier, may have some idle funds in its bank account, say $1 million, which it would like to lend for a week. Bombardier uses this excess $1 million to buy Treasury bills from a bank, which agrees to repurchase them the next week at a price slightly above Bombardier's purchase price. The net effect of this agreement is that Bombardier makes a loan of $1 million to the bank and holds $1 million of the bank's Treasury bills until the bank repurchases the bills to pay off the loan. Repurchase agreements are now an important source of bank funds. The most important lenders in this market are large corporations.

Overnight Funds

These instruments are typically overnight loans between banks of their deposits at the Bank of Canada. One reason why a bank might borrow in the overnight funds market is that it might find it does not have enough settlement balances in its deposit accounts at the Bank of Canada. It can then borrow these funds from another bank, which transfers them to the borrowing bank using the Bank of Canada's wire transfer system. This market is very sensitive to the credit needs of the banks, so the interest rate on these loans, called the **overnight interest rate**, is a closely watched barometer of the tightness of credit-market conditions in the banking system and the stance of monetary policy. When high, the overnight interest rate indicates that banks are strapped for funds; when low, it indicates that banks' credit needs are low.

Capital Market Instruments

Capital market instruments are debt and equity instruments with maturities of greater than one year. They have far wider price fluctuations than money market instruments and are considered to be fairly risky investments. The principal capital market instruments are listed in Table 2-2, which shows the amount at the end of 1990, 2000, 2010, and 2020. The Following the Financial News box discusses the capital market interest rates most frequently reported in the media.

Stocks

Stocks are equity claims on the net income and assets of a corporation. Their value of over $3 trillion in Canada, $55 trillion in the United States, and $90 trillion globally at the end of 2019 exceeds that of any other type of security in the capital market. The amount of new stock issues in any given year is typically quite small—less than 1% of the total value of shares outstanding. Individuals hold around half of the value of stocks; the rest is held by pension funds, mutual funds, and insurance companies.

Following the Financial News — Capital Market Interest Rates

The four interest rates on capital market instruments discussed most frequently in the media are

10-year Canada rate: The interest rate on Government of Canada bonds maturing in 10 years.

5-year variable-rate mortgage rate: The interest rate on a five-year, variable-rate residential mortgage.

5-year fixed-rate mortgage rate: The interest rate on a five-year, fixed-rate residential mortgage.

New-car loan rate: The interest rate on a four-year, fixed-rate new-car loan.

The data for these interest rates are reported daily in newspapers and on Internet sites such as www.bankrate.com and www.finance.yahoo.com.

TABLE 2-2 Principal Capital Market Instruments

	Amount Outstanding ($ Billions)			
Type of Instrument	**1990**	**2000**	**2010**	**2020**
Corporate stocks (market value)	112.5	243.9	378.4	667.1
Residential mortgages	245.3	429.1	976.6	1669.8
Corporate bonds	72.3	186.2	333.7	624.7
Government of Canada securities	175.9	355.9	394.3	699.9
Bank commercial loans	133.9	157.7	187.1	507.9
Consumer loans	99.6	193.8	474.3	629.8
Nonresidential mortgages	7.2	15.6	30.8	79.4

Source: Statistics Canada CANSIM series V122642, V122746, V122640, V31185443, V122631, V122707, V122656, and the authors' calculations.

Mortgages and Mortgage-Backed Securities

Mortgages are loans to households or firms to purchase land, housing, or other real structures in which the structure or land itself serves as collateral for the loans. The mortgage market is the largest debt market in Canada, with the outstanding amount of residential mortgages (used to purchase residential housing) almost 10 times the amount of commercial and farm mortgages. Mortgages are provided by financial institutions such as chartered banks, trust and loan companies, and credit unions and *caisses populaires*. However, in recent years, a growing amount of the funds for mortgages have been provided by **mortgage-backed securities**, bond-like debt instruments backed by a bundle of individual mortgages, whose interest and principal payments are collectively paid to the holders of the security. As we will see in Chapter 11, mortgage-backed securities and more complicated variants (CDOs) have become notorious because they played a key role in promoting the 2007–2009 global financial crisis. The federal government also plays an active role in the mortgage market via the Canada Mortgage and Housing Corporation (CMHC), which facilitates housing loans by providing mortgage-loan insurance.

Corporate Bonds

These long-term bonds are issued by corporations with very strong credit ratings. The typical *corporate bond* sends the holder an interest payment twice a year and pays off the face value when the bond matures. Some corporate bonds, called *convertible bonds,* have the additional feature of allowing the holder to convert them into a specified

number of shares of stock at any time up to the maturity date. This feature makes these convertible bonds more desirable to prospective purchasers than bonds without it, and it allows the corporation to reduce its interest payments because the bonds can increase in value if the price of the stock appreciates sufficiently. Because the outstanding amount of both convertible and nonconvertible bonds for any given corporation is small, corporate bonds are not nearly as liquid as other securities, such as Government of Canada bonds.

Although the size of the corporate bond market is substantially smaller than that of the stock market, with the amount of corporate bonds outstanding less than one-third that of stocks, the volume of new corporate bonds issued each year is substantially greater than the volume of new stock issues. Thus the behaviour of the corporate bond market is probably far more important to a firm's financing decisions than is the behaviour of the stock market. The principal buyers of corporate bonds are life insurance companies; pension funds and households are other large holders.

Government of Canada Bonds These intermediate-term bonds (those with initial maturities from 1 to 10 years) and long-term bonds (those with initial maturities greater than 10 years) are issued to finance the deficits of the federal government. Because they are the most widely traded bonds in Canada, they are the most liquid security traded in the capital market. They are held by the Bank of Canada, banks, households, and foreigners. These debt instruments are issued in either bearer or registered form and in denominations of $1000, $5000, $25 000, $100 000, and $1 million. In the case of **registered bonds**, the name of the owner appears on the bond certificate and is also recorded at the Bank of Canada. Some issues have the additional **call** (or **redemption**) feature of allowing them to be "called" on specified notice (usually 30 to 60 days).

Provincial and Municipal Government Bonds Provincial and municipal governments also issue bonds to finance expenditures on schools, roads, and other large programs. The securities issued by provincial governments are referred to as **provincial bonds** or **provincials** and those issued by municipal governments as **municipal bonds** or **municipals**—the securities issued by the federal government are referred to as **Canadas**. Provincials and municipals are denominated in either domestic currency or foreign currencies, mostly U.S. dollars, Swiss francs, and Japanese yen. Provincial and municipal government bonds are mainly held by trusteed pension plans, social security funds (predominantly the Canada Pension Plan), and foreigners.

Government Agency Securities These are long-term bonds issued by various government agencies, such as the Ontario Municipal Improvement Corporation and the Alberta Municipal Financing Corporation, to assist municipalities to finance such items as mortgages, farm loans, or power-generating equipment. The provincial governments guarantee many of these securities. They function much like Canadas, provincials, and municipals and are held by similar parties.

Consumer and Bank Commercial Loans These loans to consumers and businesses are made principally by banks, but in the case of consumer loans, also by finance companies.

2.4 INTERNATIONALIZATION OF FINANCIAL MARKETS

LO 2.4 Recognize the international dimensions of financial markets.

The growing internationalization of financial markets has become an important trend. Before the 1980s, U.S. financial markets were much larger than those outside the United States, but in recent years the dominance of U.S. markets has been weakening. (See the

Global Are U.S. Capital Markets Losing Their Edge?

Over the past few decades, the United States lost its international dominance in a number of manufacturing industries, including automobiles and consumer electronics, as other countries became more competitive in global markets. Recent evidence suggests that financial markets now are undergoing a similar trend: Just as Ford and General Motors have lost global market share to Toyota and Honda, recently U.S. stock and bond markets have seen their share of sales of newly issued corporate securities slip. The London and Hong Kong stock exchanges now handle a larger share of initial public offerings (IPOs) of stock than does the New York Stock Exchange, which had been by far the dominant exchange in terms of IPO value before 2000. Furthermore, the number of stocks listed on U.S. exchanges has been falling, while stock listings abroad have been growing rapidly: Listings outside the United States are now about 10 times greater than those in the United States. Likewise, the portion of new corporate bonds issued worldwide that are initially sold in U.S. capital markets has fallen below the share sold in European debt markets.

Why do corporations that issue new securities to raise capital now conduct more of this business in financial markets in Europe and Asia? Among the factors contributing to this trend are quicker adoption of technological innovation by foreign financial markets, tighter immigration controls in the United States following the terrorist attacks of 2001, and perceptions that listing on American exchanges will expose foreign securities issuers to greater risks of lawsuits.

Many people see burdensome financial regulation as the main cause, however, and point specifically to the Sarbanes-Oxley Act of 2002. The U.S. Congress passed this act after a number of accounting scandals involving U.S. corporations and the accounting firms that audited them came to light. Sarbanes-Oxley aims to strengthen the integrity of the auditing process and the quality of information provided in corporate financial statements. The costs to corporations of complying with the new rules and procedures are high, especially for smaller firms, but largely avoidable if firms choose to issue their securities in financial markets outside the United States. For this reason, there is much support for revising Sarbanes-Oxley to lessen its allegedly harmful effects and induce more securities issuers to return to U.S. financial markets. However, evidence is not conclusive to support the view that Sarbanes-Oxley is the main cause of the relative decline of U.S. financial markets and therefore in need of reform.

Discussion of the relative decline of U.S. financial markets and debate about the factors that are contributing to it likely will continue. Chapter 8 provides more detail on the Sarbanes-Oxley Act and its effects on the U.S. financial system.

Global box "Are U.S. Capital Markets Losing Their Edge?") The extraordinary growth of foreign financial markets has been the result of both large increases in the pool of savings in foreign countries such as Japan and the deregulation of foreign financial markets, which has enabled foreign markets to expand their activities. Canadian corporations and banks are now more likely to tap international capital markets to raise needed funds, and Canadian investors often seek investment opportunities abroad. Similarly, foreign corporations and banks raise funds from Canadians, and foreigners have become important investors in Canada. A look at international bond markets and world stock markets will give us a picture of how this globalization of financial markets is taking place.

International Bond Market, Eurobonds, and Eurocurrencies

The traditional instruments in the international bond market are known as **foreign bonds**. Foreign bonds are sold in a foreign country and are denominated in that country's currency. For example, if the German automaker Porsche sells a bond in Canada denominated in Canadian dollars, it is classified as a foreign bond. Foreign bonds have

been an important instrument in the international capital market for centuries. In fact, a large percentage of U.S. railroads built in the nineteenth century were financed by sales of foreign bonds in Britain.

A more recent innovation in the international bond market is the **Eurobond**, a bond denominated in a currency other than that of the country in which it is sold—for example, a bond denominated in Canadian dollars sold in London. Currently, over 80% of the new issues in the international bond market are Eurobonds, and the market for these securities has grown very rapidly. As a result, the Eurobond market is now larger than the U.S. corporate bond market.

A variant of the Eurobond is **Eurocurrencies**, which are foreign currencies deposited in banks outside the home country. The most important of the Eurocurrencies are **Eurodollars**, which are U.S. dollars deposited in foreign banks outside the United States or in foreign branches of U.S. banks. Because these short-term deposits earn interest, they are similar to short-term Eurobonds. Canadian banks borrow Eurodollar deposits from other banks or from their own foreign branches, and Eurodollars are now an important source of funds for Canadian banks.

Note that the currency, the euro, can create some confusion about the terms *Eurobond*, *Eurocurrencies*, and *Eurodollars*. A bond denominated in euros is called a Eurobond *only if it is sold outside the countries that have adopted the euro.* In fact, most Eurobonds are not denominated in euros but are instead denominated in U.S. dollars. Similarly, Eurodollars have nothing to do with euros but are instead U.S. dollars deposited in banks outside the United States.

World Stock Markets

Until recently, the U.S. stock market was by far the largest in the world, but stock markets in other countries have been growing in importance, with the United States not always number one. The increased interest in foreign stocks has prompted the development in Canada of mutual funds that specialize in trading in foreign stock markets. As the Following the Financial News box indicates, Canadian investors now pay attention not only to the Canadian stock market but also to stock price indexes for foreign stock markets such as the Dow Jones Industrial Average (New York), the Nikkei 300 Average (Tokyo), and the Financial Times Stock Exchange (FTSE) 100-Share Index (London).

The internationalization of financial markets is having profound effects on Canada. Foreigners are not only providing funds to corporations in Canada but are also helping finance the federal government. Without these foreign funds, the Canadian economy would have grown far less rapidly in the past 20 years. The internationalization of financial markets is also leading the way to a more integrated world economy in which flows of goods and technology between countries are more commonplace. In later chapters, we will encounter many examples of the important roles that international factors play in our economy.

2.5 FUNCTION OF FINANCIAL INTERMEDIARIES: INDIRECT FINANCE

LO 2.5 Summarize the roles of transaction costs, risk sharing, and information costs as they relate to financial intermediaries.

As shown in Figure 2-1 (p. 23), funds can move from lenders to borrowers by a second route, called *indirect finance* because it involves a financial intermediary that stands between the lender-savers and the borrower-spenders and helps transfer funds from

Following the Financial News — Foreign Stock Market Indexes

Foreign stock market indexes are published daily in newspapers and on Internet sites such as finance.yahoo.com.

The most important of these stock market indexes are as follows:

Dow Jones Industrial Average (DJIA): An index of the 30 largest publicly traded corporations in the United States, maintained by the Dow Jones Corporation

S&P 500: An index of 500 of the largest companies traded in the United States, maintained by Standard & Poor's

NASDAQ Composite: An index for all the stocks that trade on the NASDAQ stock market, where most of the technology stocks in the United States are traded

FTSE 100: An index of the 100 most highly capitalized UK companies listed on the London Stock Exchange

DAX: An index of the 30 largest German companies trading on the Frankfurt Stock Exchange

CAC 40: An index of the 40 largest French companies trading on Euronext Paris

Hang Seng: An index of the largest companies trading on the Hong Kong stock markets

Strait Times: An index of the 30 largest companies trading on the Singapore Exchange

These indexes are reported daily in newspapers and on Internet sites such as www.finance.yahoo.com.

one to the other. A financial intermediary does this by borrowing funds from lender-savers and then using these funds to make loans to borrower-spenders. For example, a bank might acquire funds by issuing a liability to the public in the form of savings deposits (an asset for the public). It might then use the funds to acquire an asset by making a loan to General Motors or by buying a Canada bond in the financial market. The ultimate result is that funds have been transferred from the public (the lender-savers) to General Motors or the Canadian government (the borrower-spender) with the help of the financial intermediary (the bank).

The process of indirect financing using financial intermediaries, called **financial intermediation**, is the primary route for moving funds from lenders to borrowers. Indeed, although the media focus much of their attention on securities markets, particularly the stock market, financial intermediaries are a far more important source of financing for corporations than securities markets are. This is true not only for Canada and the United States but for other industrialized countries as well (see the Global box "The Importance of Financial Intermediaries Relative to Securities Markets: An International Comparison"). Why are financial intermediaries and indirect finance so important in financial markets? To answer this question, we need to understand the roles of transaction costs, risk sharing, and information costs in financial markets.

Transaction Costs

Transaction costs, the time and money spent in carrying out financial transactions, are a major problem for people who have excess funds to lend. As we have seen, Carl the Carpenter needs $1000 for his new tool, and you know that it is an excellent investment opportunity. You have the cash and would like to lend him the money, but to protect your investment, you have to hire a lawyer to write up the loan contract that specifies how much interest Carl will pay you, when he will make these interest payments, and when he will repay you the $1000. Obtaining the contract will cost you

Global **The Importance of Financial Intermediaries Relative to Securities Markets: An International Comparison**

Patterns of financing corporations differ across countries, but one key fact emerges: Studies of the major developed countries, including the United States, Canada, the United Kingdom, Japan, Italy, Germany, and France, show that when businesses go looking for funds to finance their activities, they usually obtain them indirectly through financial intermediaries and not directly from securities markets.* Even in the United States and Canada, which have the most developed securities markets in the world, loans from financial intermediaries are far more important for corporate finance than securities markets are. The countries that have made the least use of securities markets are Germany and Japan; in these two countries, financing from financial intermediaries has been almost 10 times greater than that from securities markets. However, after the deregulation of Japanese securities markets in recent years, the share of corporate financing by financial intermediaries has been declining relative to the use of securities markets.

Although the dominance of financial intermediaries over securities markets is clear in all countries, the relative importance of bond versus stock markets differs widely across countries. In the United States, the bond market is far more important as a source of corporate finance; on average, the amount of new financing raised using bonds is 10 times the amount raised using stocks. By contrast, countries such as France and Italy make more use of equities markets than of the bond market to raise capital.

*See, for example, Colin Mayer, "Financial Systems, Corporate Finance, and Economic Development," in *Asymmetric Information, Corporate Finance, and Investment*, ed. R. Glenn Hubbard (Chicago: University of Chicago Press, 1990), 307–332.

$500. When you figure in this transaction cost for making the loan, you realize that you can't earn enough from the deal (you spend $500 to make perhaps $100) and reluctantly tell Carl that he will have to look elsewhere.

This example illustrates that small savers like you or potential borrowers like Carl might be frozen out of financial markets and thus be unable to benefit from them. Can anyone come to the rescue? Financial intermediaries can.

Financial intermediaries can substantially reduce transaction costs because they have developed expertise in lowering them and because their large size allows them to take advantage of **economies of scale**, the reduction in transaction costs per dollar of transactions as the size (scale) of transactions increases. For example, a bank knows how to find a good lawyer to produce an airtight loan contract, and this contract can be used over and over again in its loan transactions, thus lowering the legal cost per transaction. Instead of a loan contract (which may not be all that well written) costing $500, a bank can hire a top-flight lawyer for $5000 to draw up an airtight loan contract that can be used for 2000 loans at a cost of $2.50 per loan. At a cost of $2.50 per loan, it now becomes profitable for the financial intermediary to lend Carl the $1000.

Because financial intermediaries are able to reduce transaction costs substantially, they make it possible for you to provide funds indirectly to people like Carl with productive investment opportunities. In addition, a financial intermediary's low transaction costs mean that it can provide its customers with **liquidity services**, services that make it easier for customers to conduct transactions. For example, banks provide depositors with chequing accounts that enable them to pay their bills easily. In addition, depositors can earn interest on chequing and savings accounts and yet still convert them into goods and services whenever necessary.

Risk Sharing

Another benefit made possible by the low transaction costs of financial institutions is that these institutions can help reduce the exposure of investors to **risk**—that is, uncertainty about the returns investors will earn on assets. Financial intermediaries do this through the process known as **risk sharing**: They create and sell assets with risk characteristics that people are comfortable with, and the intermediaries then use the funds they acquire by selling these assets to purchase other assets that may have far more risk. Low transaction costs allow financial intermediaries to share risk at low cost, enabling them to earn a profit on the spread between the returns they earn on risky assets and the payments they make on the assets they have sold. This process of risk sharing is also sometimes referred to as **asset transformation**, because in a sense, risky assets are turned into safer assets for investors.

Financial intermediaries also promote risk sharing by helping individuals to diversify and thereby lower the amount of risk to which they are exposed. **Diversification** entails investing in a collection (**portfolio**) of assets whose returns do not always move together, with the result that overall risk is lower than for individual assets. (Diversification is just another name for the old adage "You shouldn't put all your eggs in one basket.") Low transaction costs allow financial intermediaries to pool a collection of assets into a new asset and then sell it to individuals.

Asymmetric Information: Adverse Selection and Moral Hazard

The presence of transaction costs in financial markets explains, in part, why financial intermediaries and indirect finance play such an important role in financial markets. An additional reason is that in financial markets, one party often does not know enough about the other party to make accurate decisions. This inequality is called **asymmetric information**. For example, a borrower who takes out a loan usually has better information about the potential returns and risks associated with the investment projects for which the funds are earmarked than the lender does. Lack of information creates problems in the financial system on two fronts: before the transaction is entered into and afterward.[1]

Adverse selection is the problem created by asymmetric information *before* the transaction occurs. Adverse selection occurs when one party to a transaction has information about a hidden characteristic and takes economic advantage of this information by making an agreement (transaction) with less informed parties. In financial markets, adverse selection occurs when the potential borrowers who are the most likely to produce an undesirable (*adverse*) outcome—the bad credit risks—actively seek out a loan and are thus more likely to be selected. Because adverse selection makes it more likely that loans might be made to bad credit risks, lenders may decide not to make any loans, even though good credit risks exist in the marketplace.

To understand why adverse selection occurs, suppose you have two aunts to whom you might make a loan—Aunt Louise and Aunt Sheila. Aunt Louise is a conservative type who borrows only when she has an investment she is quite sure will pay off. Aunt Sheila, by contrast, is an inveterate gambler who has just come across a get-rich-quick scheme that will make her a millionaire if she can just borrow $1000 to invest in it. Unfortunately, as with most get-rich-quick schemes, the probability is high that the investment won't pay off and that Aunt Sheila will lose the $1000.

[1]Asymmetric information and the adverse selection and moral hazard concepts are also crucial problems for the insurance industry.

Which of your aunts is more likely to call you to ask for a loan? Aunt Sheila, of course, because she has so much to gain if the investment pays off. You, however, would not want to make a loan to her because the probability is high that her investment will turn sour and she will be unable to pay you back.

If you know both your aunts very well—that is, if your information is not asymmetric—you won't have a problem, because you will know that Aunt Sheila is a bad risk and so you will not lend to her. Suppose, though, that you don't know your aunts well. You will be more likely to lend to Aunt Sheila than to Aunt Louise because Aunt Sheila will be hounding you for the loan. Because of the possibility of adverse selection, you might decide not to lend to either of your aunts, even though there are times when Aunt Louise, who is an excellent credit risk, might need a loan for a worthwhile investment.

Moral hazard is the problem created by asymmetric information *after* the transaction occurs. Moral hazard occurs when an informed party takes a hidden (unobserved) action that harms the less informed party. In financial markets, moral hazard is the risk (*hazard*) that the borrower might engage in hidden actions that are undesirable (*immoral*) from the lender's point of view, because they make it less likely that the loan will be paid back. Because moral hazard lowers the probability that the loan will be repaid, lenders may decide that they would rather not make a loan.

As an example of moral hazard, suppose that you made a $1000 loan to another relative, Uncle Melvin, who tells you he needs the money to purchase a computer so that he can set up a business typing students' term papers. Once you have made the loan, however, Uncle Melvin is more likely to secretly slip off to the track and play the horses than to purchase the computer. If he bets on a 20-to-1 long shot and wins with your money, he is able to pay back your $1000 and live high off the hog with the remaining $19 000. But if he loses, as is likely, you won't get paid back, and all he has lost is his reputation as a reliable, upstanding uncle. Uncle Melvin therefore has an incentive to go to the track because his gains ($19 000) if he bets correctly are much greater than the cost to him (his reputation) if he bets incorrectly. If you knew what Uncle Melvin was up to, you would prevent him from going to the track, and he would not be able to increase the moral hazard. However, because it is hard for you to keep informed of his whereabouts—that is, because information is asymmetric—there is a good chance that Uncle Melvin will go to the track and you will not get paid back. The risk of moral hazard might therefore discourage you from making the $1000 loan to Uncle Melvin, even if you are sure that you will be paid back if he uses it to set up his business.

The problems created by adverse selection and moral hazard are a major impediment to well-functioning financial markets. Again, financial intermediaries can alleviate these problems.

With financial intermediaries in the economy, small savers can provide their funds to the financial markets by lending these funds to a trustworthy intermediary—say, the Honest John Bank—which in turn lends the funds out either by making loans or by buying securities such as stocks or bonds. Successful financial intermediaries have higher earnings on their investments than do small savers because they are better equipped than individuals to screen out bad credit risks from good ones, thereby reducing losses due to adverse selection. In addition, financial intermediaries have high earnings because they develop expertise in monitoring the parties they lend to, thus reducing losses due to moral hazard. The result is that financial intermediaries can afford to pay lender-savers interest or provide substantial services and still earn a profit.

As we have seen, financial intermediaries play an important role in the economy because they provide liquidity services, promote risk sharing, and solve information problems, thereby allowing small savers and borrowers to benefit from the existence of financial markets. The success of financial intermediaries in performing this role is

evidenced by the fact that most Americans invest their savings with them and obtain loans from them. Financial intermediaries play a key role in improving economic efficiency because they help financial markets channel funds from lender-savers to people with productive investment opportunities. Without a well-functioning set of financial intermediaries, it is very hard for an economy to reach its full potential. We will explore further the role of financial intermediaries in the economy in Part 3.

Economies of Scope and Conflicts of Interest

Another reason why financial intermediaries play such an important part in the economy is that by providing multiple financial services to their customers, such as offering them bank loans or selling their bonds for them, they can also achieve **economies of scope**; that is, they can lower the cost of information production for each service by applying one information resource to many different services. A bank, for example, when making a loan to a corporation, can evaluate how good a credit risk the firm is, which then helps the bank decide whether it would be easy to sell the bonds of this corporation to the public.

Although economies of scope may substantially benefit financial institutions, they also create potential costs in terms of **conflicts of interest**. Conflicts of interest, a type of moral hazard problem, arise when a person or institution has multiple objectives (interests), some of which conflict with each other. Conflicts of interest are especially likely to occur when a financial institution provides multiple services. The potentially competing interests of those services may lead an individual or firm to conceal information or disseminate misleading information. We care about conflicts of interest because a substantial reduction in the quality of information in financial markets increases asymmetric information problems and prevents financial markets from channelling funds into the most productive investment opportunities. Consequently, the financial markets and the economy become less efficient.

2.6 TYPES OF FINANCIAL INTERMEDIARIES

LO 2.6 List and describe the different types of financial intermediaries.

We have seen why financial intermediaries have such an important function in the economy. Now we take a look at the principal financial intermediaries and how they perform the intermediation function. Financial intermediaries fall into three categories: depository institutions (banks), contractual savings institutions, and investment intermediaries. Table 2-3 describes the primary liabilities (sources of funds) and assets (uses of funds) of the financial intermediaries in each category. The relative sizes of these intermediaries in Canada are indicated in Table 2-4.

Depository Institutions

Depository institutions (for simplicity, we refer to these as *banks* throughout this text) are financial intermediaries that accept deposits from individuals and institutions and make loans. The study of money and banking focuses special attention on this group of financial institutions because they are involved in the creation of deposits, an important component of the money supply. These institutions include commercial banks

TABLE 2-3 Primary Assets and Liabilities of Financial Intermediaries

Type of Intermediary	Primary Liabilities (Sources of Funds)	Primary Assets (Uses of Funds)
Depository institutions (banks)		
Chartered banks	Deposits	Loans, mortgages, government bonds
Trust and loan companies	Deposits	Mortgages
Credit unions and *caisses populaires*	Deposits	Mortgages
Contractual savings institutions		
Life insurance companies	Premiums from policies	Corporate bonds and mortgages
Property and casualty insurance companies	Premiums from policies	Corporate bonds and stocks
Pension funds	Retirement contributions	Corporate bonds and stocks
Investment intermediaries		
Finance companies	Finance paper, stocks, bonds	Consumer and business loans
Mutual funds	Shares	Stocks and bonds
Money market mutual funds	Shares	Money market instruments

and the so-called near banks: trust and loan companies, credit unions, and *caisses populaires*.

Chartered Banks These financial intermediaries raise funds primarily by issuing chequable deposits (deposits on which cheques can be written), savings deposits (deposits that are payable on demand), and term deposits (deposits with fixed terms to maturity). They then use these funds to make commercial, consumer, and mortgage loans and to buy Canadian government securities and provincial and municipal bonds. Around 80 chartered banks are found in Canada, and as a group, they are the largest financial intermediary and have the most diversified portfolios (collections) of assets.

Trust and Loan Companies These depository institutions obtain funds primarily through chequable and nonchequable savings deposits, term deposits, guaranteed investment certificates, and debentures. In the past, these institutions were constrained in their activities and mostly made mortgage loans for residential housing. Over time, these restrictions have been loosened so that the distinction between these depository institutions and chartered banks has blurred. These intermediaries have become more alike and are now more competitive with each other.

Credit Unions and *Caisses Populaires* (CUCPs) Credit unions and *caisses populaires* are very small cooperative lending institutions organized around a particular group: union members, employees of a particular firm, and so forth. They acquire funds from deposits and primarily make mortgage and consumer loans.

TABLE 2-4 Relative Shares of Financial Institutions and Pension Plans Regulated by OSFI (as of December 31, 2020)

Type of Intermediary	Total Assets ($ Millions)	Percent (%)
Chartered Banks		
Domestic	6 909 661	60.34
Foreign bank subsidiaries	77	0.00
Foreign bank branches	1876	0.02
Trust Companies	155 988	1.36
Loan Companies	249 636	2.18
Cooperative Credit Associations	270 200	2.36
Life Insurance Companies		
Canadian-incorporated	1 683 840	14.71
Foreign branches	19 116	0.17
Fraternal Benefit Societies		
Canadian-incorporated	18 770	0.16
Foreign branches	3944	0.03
Property and Casualty Insurance Companies		
Canadian-incorporated	132 000	1.15
Foreign branches	35 569	0.31
Pension Plans	1 970 000	17.20
Total	11 450 677	100.00

Source: Statistics Canada and OSFI websites.

Contractual Savings Institutions

Contractual savings institutions, such as insurance companies and pension funds, are financial intermediaries that acquire funds at periodic intervals on a contractual basis. Because they can predict with reasonable accuracy how much they will have to pay out in benefits in the coming years, they do not have to worry as much as depository institutions about losing funds quickly. As a result, the liquidity of assets is not as important a consideration for them as it is for depository institutions, and they tend to invest their funds primarily in long-term securities such as corporate bonds, stocks, and mortgages.

Life Insurance Companies Life insurance companies insure people against financial hazards following a death and sell annuities (annual income payments upon retirement). They acquire funds from the premiums that people pay to keep their policies in force and use them mainly to buy corporate bonds and mortgages. They also purchase stocks but are restricted in the amount that they can hold. Currently, with

close to $1.7 trillion of assets in Canada (and over US$8 trillion of assets in the United States), they are among the largest of the contractual savings institutions.

Property and Casualty (P&C) Insurance Companies These companies insure their policyholders against loss from theft, fire, and accidents. They are very much like life insurance companies, receiving funds through premiums for their policies, but they have a greater possibility of loss of funds if major disasters occur. For this reason, they use their funds to buy more liquid assets than life insurance companies do. Their largest holding of assets is government bonds and debentures; they also hold corporate bonds and stocks.

Pension Funds and Government Retirement Funds Private pension funds and provincial and municipal retirement funds provide retirement income in the form of annuities to employees who are covered by a pension plan. Funds are acquired by contributions from employers and from employees, who either have a contribution automatically deducted from their paycheque or contribute voluntarily. The largest asset holdings of pension funds are corporate bonds and stocks. The establishment of pension funds has been actively encouraged by the federal government, both through legislation requiring pension plans and through tax incentives to encourage contributions.

Investment Intermediaries

This category of financial intermediary includes finance companies, mutual funds, money market mutual funds, and hedge funds.

Finance Companies Finance companies raise funds by selling commercial paper (a short-term debt instrument) and by issuing stocks and bonds. They lend these funds to consumers, who use them to purchase such items as furniture, automobiles, and home improvements, and to small businesses. Some finance companies are organized by a parent corporation to help sell its product. For example, Ford Motor Credit Company makes loans to consumers who purchase Ford automobiles.

Mutual Funds These financial intermediaries acquire funds by selling shares to many individuals and then using the proceeds to purchase diversified portfolios of stocks and bonds. Mutual funds allow shareholders to pool their resources so that they can take advantage of lower transaction costs when buying large blocks of stocks or bonds. In addition, mutual funds allow shareholders to hold more diversified portfolios than they otherwise would. Shareholders can sell (redeem) shares at any time, but the value of these shares will be determined by the value of the mutual fund's holdings of securities. Because these fluctuate greatly, the value of mutual fund shares does too; therefore, investments in mutual funds can be risky.

Money Market Mutual Funds These financial institutions are similar to mutual funds but also function to some extent as depository institutions because they offer deposit-type accounts. Like most mutual funds, they sell shares to acquire funds. These funds are then used to buy money market instruments that are both safe and very liquid. The interest on these assets is paid out to the shareholders.

Hedge Funds Hedge funds are a type of mutual fund with special characteristics. Hedge funds are organized as limited partnerships with minimum investments ranging from $100 000 to, more typically, $1 million or more. These limitations mean that hedge funds are subject to much weaker regulation than other mutual funds. Hedge funds invest in many types of assets, with some specializing in stocks, others in bonds, others in foreign currencies, and still others in far more exotic assets.

Investment Banks Despite its name, an investment bank is not a bank or a financial intermediary in the ordinary sense; that is, it does not take in funds and then lend them out. Instead, an investment bank is a different type of intermediary that helps a corporation issue securities. First it advises the corporation on which type of securities to issue (stocks or bonds); then it helps sell (**underwrite**) the securities by purchasing them from the corporation at a predetermined price and reselling them in the market. Investment banks also act as deal makers and earn enormous fees by helping corporations acquire other companies through mergers or acquisitions.

2.7 REGULATION OF THE FINANCIAL SYSTEM

LO 2.7 Identify the reasons for and list the types of financial market regulations.

The financial system is among the most heavily regulated sectors of the Canadian economy. The government regulates financial markets for two main reasons: to increase the information available to investors and to ensure the soundness of the financial system. We will examine how these two goals have led to the present regulatory environment. As a study aid, the principal regulatory agencies of the Canadian financial system are listed in Table 2-5.

Increasing Information Available to Investors

Asymmetric information in financial markets means that investors may be subject to adverse selection and moral hazard problems that may hinder the efficient operation of financial markets. Risky firms or outright crooks may be the most eager to sell securities to unwary investors, and the resulting adverse selection problem may keep investors out of financial markets. Furthermore, once an investor has bought a security, thereby lending money to a firm, the borrower may have incentives to engage in risky activities or to commit outright fraud. The presence of this moral hazard problem may also keep investors away from financial markets. Government regulation can reduce adverse selection and moral hazard problems in financial markets and enhance the efficiency of the markets by increasing the amount of information available to investors.

Provincial securities commissions, the most significant being the Ontario Securities Commission (OSC), administer provincial acts requiring corporations issuing securities to disclose certain information about their sales, assets, and earnings to the public and restrict trading by the largest stockholders in the corporation. By requiring disclosure of this information and by discouraging insider trading, which could be used to manipulate security prices, regulators hope that investors will be better informed and be protected from abuses in financial markets. Indeed, in recent years the OSC has been particularly active in prosecuting people involved in insider trading in Canada's largest stock exchange, the Toronto Stock Exchange.

TABLE 2-5 Principal Regulatory Agencies of the Canadian Financial System

Regulatory Agency	Subject of Regulation	Nature of Regulations
Provincial securities and exchange commissions	Organized exchanges and financial markets	Requires disclosure of information and restricts insider trading
Bank of Canada	Chartered banks, TMLs, and CUCPs	Examines the books of the deposit-taking institutions and coordinates with the federal agencies that are responsible for financial institution regulation: OSFI and CDIC
Office of the Superintendent of Financial Institutions Canada (OSFI)	All federally regulated chartered banks, TMLs, CUCPs, life insurance companies, P&C insurance companies, and pension plans	Sets capital adequacy, accounting, and board-of-directors responsibility standards; conducts bank audits and coordinates with provincial securities commissions
Canada Deposit Insurance Corporation (CDIC)	Chartered banks, TMLs, CUCPs	Provides insurance of up to $100 000 for each depositor at a bank, examines the books of insured banks, and imposes restrictions on assets they can hold
Québec Deposit Insurance Board	TMLs and credit cooperatives in Québec	Similar role to the CDIC
Canadian Life and Health Insurance Compensation Corporation (CompCorp)	Life insurance companies	Compensates policyholders if the issuing life insurance company goes bankrupt
P&C Insurance Compensation Corporation (PACIC)	Property and casualty insurance companies	Compensates policyholders if the issuing P&C insurance company goes bankrupt

Ensuring the Soundness of Financial Intermediaries

Asymmetric information can lead to the widespread collapse of financial intermediaries, referred to as a **financial panic**. Because providers of funds to financial intermediaries may not be able to assess whether the institutions holding their funds are sound, if they have doubts about the overall health of financial intermediaries, they may want to pull their funds out of both sound and unsound institutions. The possible outcome is a financial panic that produces large losses for the public and causes serious damage to the economy. To protect the public and the economy from financial panics, the government has implemented six types of regulations.

Restrictions on Entry Provincial banking and insurance commissions, the Bank of Canada, and the Office of the Superintendent of Financial Institutions (OSFI), an agency of the federal government, have created tight regulations governing who is allowed to set up a financial intermediary. Individuals or groups that want to establish

a financial intermediary, such as a bank or an insurance company, must obtain a charter from the provincial or federal government. Only upstanding citizens with impeccable credentials and a large amount of initial funds will be given a charter.

Disclosure Reporting requirements for financial intermediaries are stringent. Their bookkeeping must follow certain strict principles, their books are subject to periodic inspection, and they must make certain information available to the public.

Restrictions on Assets and Activities Financial intermediaries are restricted in what they are allowed to do and what assets they can hold. Before you put your funds into a chartered bank or similar institution, you will want to know that your funds are safe and that the bank or other financial intermediary will be able to meet its obligations to you. One way of ensuring the trustworthiness of financial intermediaries is to restrict them from engaging in certain risky activities. Another way to limit a financial intermediary's behaviour is to restrict it from holding certain risky assets, or at least from holding a greater quantity of these risky assets than is prudent. For example, chartered banks and other depository institutions are not allowed to hold common stock because stock prices experience substantial fluctuations. Insurance companies are allowed to hold common stock, but their holdings cannot exceed a certain fraction of their total assets.

Deposit Insurance The government can insure people's deposits so that they do not suffer any financial loss if the financial intermediary that holds these deposits should fail. The most important government agency that provides this type of insurance is the Canada Deposit Insurance Corporation (CDIC), created by an act of Parliament in 1967. It insures each depositor at a member deposit-taking financial institution up to a loss of $100 000 per account. Except for certain wholesale branches of foreign banks, credit unions, and some provincial institutions, all deposit-taking financial institutions in Canada are members of the CDIC. All CDIC members make contributions into the CDIC fund, which are used to pay off depositors in the case of a bank's failure. The Québec Deposit Insurance Board, an organization similar to CDIC and set up at the same time as CDIC, provides insurance for TMLs and credit cooperatives in Québec.

Limits on Competition Politicians have often declared that unbridled competition among financial intermediaries promotes failures that harm the public. Although the evidence that competition has this effect is extremely weak, provincial and federal governments at times have imposed restrictive regulations. For example, from 1967 to 1980 the entry of foreign banks into Canadian banking was prohibited. Since 1980, however, the incorporation of foreign bank subsidiaries has been regulated, but Canada still ranks low with respect to the degree of competition from foreign banks.

In later chapters we will look more closely at government regulation of financial markets and will see whether it has improved their functioning.

Financial Regulation Abroad

Not surprisingly, given the similarity of the economic systems here and in the United States, Japan, and the nations of Western Europe, financial regulation in these countries is similar to financial regulation in Canada. Provision of information is improved by requiring corporations issuing securities to report details about assets and liabilities, earnings, and sales of stock, and by prohibiting insider trading. The soundness of

intermediaries is ensured by licensing, periodic inspection of financial intermediaries' books, and the provision of deposit insurance.

The major differences between financial regulation in Canada and abroad relate to bank regulation. In the past, for example, the United States was the only industrialized country to subject banks to restrictions on branching, which limited their size and confined them to certain geographic regions. (These restrictions were abolished by legislation in 1994.) U.S. and Canadian banks are also the most restricted in the range of assets they may hold. Banks in other countries frequently hold shares in commercial firms; in Japan and Germany, those stakes can be sizable.

SUMMARY

1. The basic function of financial markets is to channel funds from savers who have an excess of funds to spenders who have a shortage of funds. Financial markets can do this either through direct finance, in which borrowers borrow funds directly from lenders by selling them securities, or through indirect finance, which involves a financial intermediary that stands between the lender-savers and the borrower-spenders and helps transfer funds from one to the other. This channelling of funds improves the economic welfare of everyone in society. Because they allow funds to move from people who have no productive investment opportunities to those who have such opportunities, financial markets contribute to economic efficiency. In addition, channelling of funds directly benefits consumers by allowing them to make purchases when they need them most.
2. Financial markets can be classified as debt and equity markets, primary and secondary markets, exchanges and over-the-counter markets, and money and capital markets.
3. The principal money market instruments (debt instruments with maturities of less than one year) are Treasury bills, negotiable bank certificates of deposit, commercial paper, repurchase agreements, and overnight funds. The principal capital market instruments (debt and equity instruments with maturities greater than one year) are stocks, mortgages, corporate bonds, Canadian government securities, Canada Savings Bonds, government agency securities, provincial and municipal government bonds, and consumer and bank commercial loans.
4. An important trend in recent years is the growing internationalization of financial markets. Eurobonds, which are denominated in a currency other than that of the country in which they are sold, are now the dominant security in the international bond market. Eurodollars, which are dollars deposited in foreign banks, are an important source of funds for Canadian banks.
5. Financial intermediaries are financial institutions that acquire funds by issuing liabilities and, in turn, use those funds to acquire assets by purchasing securities or making loans. Financial intermediaries play an important role in the financial system because they reduce transaction costs, allow risk sharing, and solve problems created by adverse selection and moral hazard. As a result, financial intermediaries allow small savers and borrowers to benefit from the existence of financial markets, thereby increasing the efficiency of the economy. However, the economies of scope that help make financial intermediaries successful can lead to conflicts of interest that make the financial system less efficient.
6. The principal financial intermediaries fall into three categories: (a) banks—chartered banks, trust and loan companies, and credit unions and *caisses populaires*; (b) contractual savings institutions—life insurance companies, fire and casualty insurance companies, and pension funds; and (c) investment intermediaries—finance companies, mutual funds, money market mutual funds, hedge funds, and investment banks.
7. The government regulates financial markets and financial intermediaries for two main reasons: to increase the information available to investors and to ensure the soundness of the financial system. Regulations include requiring disclosure of information to the public, restrictions on who can set up a financial intermediary, restrictions on what assets financial intermediaries can hold, the provision of deposit insurance, limits on competition, and restrictions on interest rates.

MyLab Economics Make the grade with MyLab Economics: Select end-of-chapter questions can be found on MyLab Economics. You can practise them as often as you want, and most feature step-by-step guided instructions to help you find the right answer.

QUESTIONS

1. If I can buy a car today for $5000 and it is worth $10 000 in extra income to me next year because it enables me to get a job as a travelling salesman, should I take out a loan from Larry the Loan Shark at a 90% interest rate if no one else will give me a loan? Will I be better or worse off as a result of taking out this loan? Can you make a case for legalizing loan sharking?
2. Some economists suspect that one of the reasons economies in developing countries grow so slowly is that they do not have well-developed financial markets. Does this argument make sense?
3. Give at least three examples of a situation in which financial markets allow consumers to better time their purchases.
4. If you suspect that a company will go bankrupt next year, which would you rather hold—bonds issued by the company or equities issued by the company? Why?
5. Suppose that Toyota sells yen-denominated bonds in Tokyo. Is this debt instrument considered a Eurobond? How would your answer change if the bond were sold in New York?
6. Describe who issues each of the following money market instruments:
   a. Treasury bills
   b. Certificates of deposit
   c. Commercial paper
   d. Repurchase agreement
   e. Overnight funds
7. What is the difference between a *mortgage* and a *mortgage-backed security*?
8. The U.S. economy borrowed heavily from the British in the nineteenth century to build a railroad system. Why did this make both countries better off?
9. A significant number of European banks held large amounts of assets as mortgage-backed securities derived from the U.S. housing market, which crashed after 2006. How does this demonstrate both a benefit and a cost to the internationalization of financial markets?
10. How does risk sharing benefit both financial intermediaries and private investors?
11. How can the adverse selection problem explain why you are more likely to make a loan to a family member than to a stranger?
12. One of the factors contributing to the financial crisis of 2007–2009 was the widespread issuance of subprime mortgages. How does this demonstrate adverse selection?
13. Why do loan sharks worry less about moral hazard in connection with their borrowers than some other lenders do?
14. If you are an employer, what kinds of moral hazard problems might you worry about with regard to your employees?
15. If there were no asymmetry in the information that a borrower and a lender had, could a moral hazard problem still exist?
16. "In a world without information costs and transaction costs, financial intermediaries would not exist." Is this statement true, false, or uncertain? Explain your answer.
17. Why might you be willing to make a loan to your neighbour by putting funds in a savings account earning a 5% interest rate at the bank and having the bank lend them the funds at a 10% interest rate rather than lend them the funds yourself?
18. How do conflicts of interest make the asymmetric information problem worse?
19. How can the provision of several types of financial services by one firm be both beneficial and problematic?
20. If you were going to get a loan to purchase a new car, which financial intermediary would you use: a credit union, a pension fund, or an investment bank?
21. Why would a life insurance company be concerned about the financial stability of major corporations or the health of the housing market?

22. In 2008, as a financial crisis began to unfold in the United States, the FDIC raised the limit on insured losses to bank depositors from $100 000 per account to $250 000 per account. How would this help stabilize the financial system?

23. Financial regulation is similar, but not exactly the same, in industrialized countries. Discuss why it might be desirable—or undesirable—to have the same financial regulation across industrialized countries.

APPLIED PROBLEMS

24. Suppose you have just inherited $10 000 and are considering the following options for investing the money to maximize your return:

Option 1: Put the money in an interest-bearing chequing account that earns 2%. The CDIC insures the account against bank failure.

Option 2: Invest the money in a corporate bond with a stated return of 5%, although there is a 10% chance the company could go bankrupt.

Option 3: Loan the money to one of your friend's roommates, Mike, at an agreed-on interest rate of 8%, even though you believe there is a 7% chance that Mike will leave town without repaying you.

Option 4: Hold the money in cash and earn zero return.

a. If you are risk-neutral (that is, neither seek out nor shy away from risk), which of the four options should you choose to maximize your expected return? (*Hint:* To calculate the *expected return* of an outcome, multiply the probability that an event will occur by the outcome of that event and then add them up.)

b. Suppose Option 3 and Option 4 are your only choices. If you could pay your friend $100 to find out extra information about Mike that would indicate with certainty whether he will leave town without paying, would you pay the $100? What does this say about the value of better information regarding risk?

DATA ANALYSIS PROBLEMS

1. Go to the CANSIM database and find data on the total Canadian dollar assets held by chartered banks in Canada (series V36852). What is the percentage increase in growth of assets from January 2020 to November 2020?

2. **Real-Time Data Analysis** Go to the St. Louis Federal Reserve FRED database and find data on the total assets of all U.S. commercial banks (TLAACBM027SBOG) and the total assets of all money market mutual funds (MMMFFAQ027S). Transform the commercial bank assets series to quarterly by adjusting the *Frequency* setting to "Quarterly." Calculate the percentage increase in growth of assets for each series, from January 2000 to the most recent quarter available. Which of the two financial intermediaries has experienced the most percentage growth?

SUPPLEMENTAL CHAPTER

2 Nonbank Finance

Learning Objectives

S2.1 Summarize the different types of insurance products and ways in which insurance companies can reduce asymmetric information problems.

S2.2 Summarize the distinctions between defined-benefit and defined-contribution pensions and the key features of private and public pension plans.

S2.3 List and describe the different types of finance companies.

S2.4 Summarize the roles of investment banks, securities brokers and dealers, and organized exchanges in the securities market.

S2.5 Describe the role and activities of mutual funds in financial intermediation.

S2.6 Summarize the key distinctions between hedge funds and mutual funds.

S2.7 Define private equity and venture capital, and summarize their advantages as investment funds.

Preview

Banking is not the only type of financial intermediation you are likely to encounter. You might decide to purchase insurance, take out an installment loan from a finance company, or buy a share of stock. In carrying out each of these transactions, you will be engaging in nonbank finance and dealing with nonbank financial institutions. In our economy, nonbank finance plays an important role in channelling funds from lender-savers to borrower-spenders. Furthermore, the process of financial innovation we discussed in Chapter 10 has increased the importance of nonbank finance and has led to the growth of the shadow banking system, which played such a prominent role in the recent global financial crisis (Chapter 12). In this chapter, we will examine in more detail how institutions engaged in nonbank finance operate and how they are regulated. We will also examine recent trends in nonbank finance.

S2.1 INSURANCE

LO S2.1 Summarize the different types of insurance products and ways in which insurance companies can reduce asymmetric information problems.

Every day we face the possibility of the occurrence of catastrophic events that might lead to large financial losses. A spouse's earnings might disappear due to death or illness; a car accident might result in costly repair bills or payments to an injured party. Because financial losses from crises could be large relative to our financial resources, we protect ourselves against them by purchasing insurance coverage that will pay a sum of money if a catastrophic event occurs. Life insurance companies sell policies that provide income if a person dies, is incapacitated by illness, or retires. Property and casualty companies specialize in policies that pay for losses incurred as a result of accidents, fire, or theft.

Life Insurance

There are currently more than 150 life and health insurance companies operating in Canada (including 38 foreign-owned), which are organized as either stock companies or mutuals. Shareholders own stock companies; mutuals are technically similar to credit unions, owned by the policyholders. Prior to 1999, half of the life insurance companies in Canada were organized as mutuals. In 1999, however, many of the large

S2.8 Describe the two types of government financial intermediation.

mutual life insurance companies started a process called **demutualization** and have now converted to stock companies.

Life insurance company regulation is the responsibility of OSFI and Assuris. OSFI regulation is primarily prudential, directed at sales practices, the provision of adequate liquid assets to cover losses, and restrictions on the amount of risky assets (such as common stock) the companies can hold. In other words, OSFI performs the same oversight functions as it does for banks and near banks. Assuris has no regulatory role in overseeing individual life insurance companies. It is a federally incorporated, private, nonprofit corporation established and funded by the Canadian life and health insurance industry to provide liability insurance to policyholders: It compensates policyholders if the issuing company goes bankrupt.

Because death rates for the population as a whole are predictable with a high degree of certainty, life insurance companies can accurately predict their future payouts to policyholders. Consequently, they hold long-term assets that are not particularly liquid—corporate bonds and commercial mortgages as well as some corporate stock.

Life insurance policies are offered in two principal forms: permanent life insurance (such as whole, universal, and variable life) and temporary life insurance (such as term). Permanent life insurance policies require the payment of a fixed premium throughout the life of the policy. In the early years of the policy, the size of this premium exceeds the cost of insuring against death because the probability of death is low. Thus the policy builds up a cash value in its early years. In later years, however, the cash value declines because the amount of the fixed premium falls below the cost of insuring against death, the probability of which is now higher. The policyholder can borrow against the cash value of the permanent life policy or can claim it by cancelling the policy.

Term insurance, by contrast, is characterized by a premium that is matched every year, for the period of the term (such as one year or five years), to the cost of insuring against death. As a result, a term policy requires the policyholder to pay a premium that increases over time as the probability of death increases (or else to pay a level premium, characterized by a decline over the term of the policy in the amount of death benefits). Term policies have no cash value and thus, in contrast to permanent life policies, provide insurance only, with no savings aspect.

Beginning in the mid-1970s, life insurance companies decided to restructure their industry and began to manage assets for pension funds. Now more than half of the assets managed by life insurance companies are for pension funds and not for life insurance. Insurance companies have also begun to sell investment vehicles for retirement, such as **annuities**, arrangements whereby the customer pays an annual premium in exchange for a future stream of annual payments beginning at a set age—say, 65—and continuing until death. As a result of this new business, the market share of life insurance companies as a percentage of total financial intermediary assets has held steady since 1980.

Property and Casualty Insurance

There are more than 198 property and casualty (P&C) or "general" insurance companies in Canada, and about 27% of the industry, measured by assets held, is foreign-controlled. Most property and casualty insurance companies in Canada are federally registered and subject to regulation by OSFI and the Property and Casualty Insurance Compensation Corporation (PACICC). PACICC was set up in 1988 and performs the same role for property and casualty companies as Assuris does for life insurance

companies. Some lines of P&C insurance, for example, auto insurance, are also subject to provincial laws and regulations.

Because of high interest rates in the 1970s and 1980s, property and casualty insurance companies had high investment income that enabled them to keep insurance rates low. Since then, however, interest rates have declined and investment income has fallen, while growth in the number of lawsuits involving property and casualty insurance, along with the explosion of the sums awarded in such cases, lowered their profits.

In order to return to profitability, property and casualty insurance companies raised their rates and have refused to provide coverage for some people. They have also campaigned actively for limits on insurance payouts, particularly on payouts for medical malpractice. In their search for profits, these types of insurance companies have branched out into uncharted territory by insuring the payment of interest on municipal and corporate bonds and on mortgage-backed securities. This led to insurance companies taking on excessive risk in an attempt to boost their profits. Concern about the health of the property and casualty insurance industry, particularly after the near death of AIG discussed later in this chapter, led insurance regulators to implement new rules that impose risk-based capital requirements on these insurance companies. The capital requirements are based on the riskiness of the companies' assets and operations.

The investment policies of property and casualty insurance companies are affected by the fact that property losses are very uncertain. In fact, because property losses are more uncertain than the death rate in a population, these insurers are less able to predict how much they will have to pay policyholders than life insurance companies are. Natural disasters such as the ice storm of 1998 and the Alberta flood of 2013 exposed the property and casualty insurance companies to billions of dollars of losses.[1] Therefore, property and casualty insurers hold more liquid assets than life insurers do: Cash, due and accrued investment income, money market instruments, and receivables amount to over a third of their assets, and most of the remainder is held in bonds, debentures, and stocks. Their largest liability relates to unpaid claims and adjustment expenses, followed by unearned premiums (premiums representing the unexpired part of policies).

Property and casualty insurance companies will insure against losses from almost any type of event, including fire, theft, negligence, malpractice, earthquakes, and automobile accidents. If a potential loss is too large for any one firm to insure against, several firms may join together to write a policy and thus share the risk. Insurance companies may also reduce their risk exposure by obtaining **reinsurance**. Reinsurance allocates a portion of the risk to another company in exchange for a portion of the premium and is particularly important for small insurance companies. You can think of reinsurance as insurance for the insurance company. The most famous risk-sharing operation is Lloyd's of London, an association in which different insurance companies can underwrite a fraction of an insurance policy. Lloyd's of London has claimed that it will insure against any contingency—for a price.

[1]For example, the freezing rains of January 1998 triggered about 840 000 insurance claims. As a result, Canadian P&C insurers paid out over $1.44 billion. Similar disasters in the United States, such as the Los Angeles earthquake in 1994; Hurricane Katrina, which devastated New Orleans in 2005; Hurricane Ike, which devastated Galveston in 2008; and the September 11, 2001, destruction of the World Trade Center, exposed the U.S. property and casualty insurance companies to billions of dollars of losses.

The Competitive Threat from the Banking Industry

Banks used to be restricted in their ability to sell life insurance products. This has changed, however. In recent years, the bank regulatory authorities have encouraged banks to enter the insurance field in order to help diversify banks' business, thereby improving their economic health and making bank failures less likely. As a result, most banks now sell insurance products of one form or another in addition to their traditional offerings, thus blurring the distinction between insurance companies and banks.

Credit Insurance

Insurance companies have not taken the encroachment of banks on their territory lying down. They have responded by getting into the business of supplying credit insurance. They have done this in two ways.

Credit Default Swaps

One way in which insurance companies can, in effect, provide credit insurance is by selling a traded derivative called a **credit default swap (CDS)**. In a credit default swap, the seller is required to make a payment to the holder of the CDS if there is a *credit event* for that instrument, such as a bankruptcy or downgrading of the firm's credit rating. (Credit derivatives are discussed more extensively in Chapter 13.) Issuing a CDS is thus tantamount to providing insurance on the debt instrument because, just like insurance, it makes a payment to the holder of the CDS when a negative credit event occurs. Major insurance companies have entered the CDS market in recent years, sometimes to their great regret (see the FYI box "The AIG Blowup").

FYI The AIG Blowup

American International Group, better known as AIG, was a trillion-dollar insurance giant and prior to 2008 was one of the 20 largest companies in the world. A small separate unit, AIG's Financial Products division, went into the credit default swap business in a big way, insuring over US$400 billion of securities, of which US$57 billion were debt securities backed by subprime mortgages. Lehman Brothers' troubles and eventual bankruptcy on September 15, 2008, revealed that subprime securities were worth much less than the value stated on the books, and investors came to the realization that AIG's losses, which had already been substantial in the first half of the year, could bankrupt the company. Lenders to AIG then pulled back with a vengeance, and AIG could not raise enough capital to stay afloat.

On September 16, 2008, the Federal Reserve and the U.S. Treasury decided to rescue AIG because its failure was deemed potentially catastrophic for the financial system. Banks and mutual funds were large holders of AIG's debt, and the bankruptcy of AIG would have rendered worthless all of the credit default swaps it had sold, thereby imposing huge losses on the financial institutions that had bought them. The Federal Reserve set up an US$85 billion credit facility (with the total loans from the Fed and the government eventually reaching US$173 billion) to provide liquidity to AIG. The rescue did not come cheaply, however: AIG was charged a very high interest rate on the loans from the Fed, and the government was given the rights to an 80% stake in the company if it survived. Maurice Greenberg, former CEO of the company, described the government's actions as a "nationalization" of AIG.

Insurance companies had never been viewed as posing a risk to the entire financial system, so their regulation had been left to insurance commissions in each state. Because the problems at AIG nearly brought down the U.S. financial system, this view is no longer tenable. The insurance industry will never be the same.

Monoline Insurance An alternative way of providing credit insurance is to supply it directly, just as is done with any insurance policy. However, insurance regulations do not allow property and casualty insurance companies, life insurance companies, or insurance companies with multiple lines of business to underwrite credit insurance. **Monoline insurance companies**, which specialize in credit insurance alone, are therefore the only insurance companies that are allowed to provide insurance that guarantees the timely repayment of bond principal and interest if a debt issuer defaults. These insurance companies have become particularly important in the municipal bond market, where they insure a large percentage of these securities. When a municipal security with a lower credit rating (say, an A rating) buys an insurance policy from a monoline insurer, it takes on the credit rating (say, AAA) of the monoline insurer. This action lowers the interest cost for the municipality and so makes it worthwhile for the municipality to pay the premium for this insurance policy. Of course, for this arrangement to work, the monoline insurer must have a very high credit rating. When monoline insurers experienced credit rating downgrades during the global financial crisis, they did not suffer alone—so did the municipal bond market (see the FYI box "The Global Financial Crisis and the Monoline Insurers").

The New Legislative Framework

As noted in Chapter 12, in 2001 the government passed legislation reforming the regulatory framework governing Canada's financial services sector. The new legislation allows demutualized life and health insurance companies to restructure under a holding company structure, to enter into joint ventures and strategic alliances, and to access the Canadian payments and clearance systems in an attempt to bring the sector in line with the banking sector. However, the new legislation does not allow mergers involving large banks and large demutualized life and health insurance companies. Moreover, the new legislation requires that large life and health insurance companies (those with equity over $5 billion) be widely held, in the sense that an individual or firm cannot own more than 20% of the voting shares. Small demutualized companies (those with equity under $1 billion) are eligible to be closely held.

FYI The Global Financial Crisis and the Monoline Insurers

One spillover from the global financial crisis was a hit to the monoline insurers in the United States, with subsequent effects on the U.S. municipal bond market. Unfortunately for the monoline insurers, they insured not only municipal bonds but also debt securities backed by subprime mortgages. As defaults on these mortgages grew, monoline insurers began to suffer big losses, resulting in credit downgrades from their AAA status. These downgrades weakened the value of their insurance guarantees, not only on subprime securities but on municipal securities as well. As the global crisis got into full swing, the markets took the view that monoline insurance was not worth much, and so municipal bonds began to trade at lower prices that were based solely on the municipality's credit rating. As a result, state and local governments now found their interest costs rising. They were hit by a double whammy of higher borrowing costs and lower tax revenues because of their weaker economies. The result was weaker state and local finances as well as cutbacks in spending on roads, schools, and hospitals.

APPLICATION Insurance Management

Insurance, like banking, is the financial intermediation business of transforming one type of asset into another for the public. Insurance providers use the premiums paid on policies to invest in assets such as bonds, stocks, mortgages, and other loans; the earnings from these assets are then used to pay out claims on the policies. In effect, insurers transform assets such as bonds, stocks, and loans into insurance policies that provide a set of services (for example, claim adjustments, savings plans, and friendly insurance agents). If the insurer's production process of asset transformation efficiently provides its customers with adequate insurance services at low cost, and if it can earn high returns on its investments, it will make profits; if not, it will suffer losses.

In Chapter 12, we applied the economic concepts of adverse selection and moral hazard to help us understand the principles of bank management as they relate to credit risk; many of these same principles apply to the lending activities of insurers. Here again we will apply the adverse selection and moral hazard concepts to explain many of the management practices that are specific to insurance.

In the case of an insurance policy, moral hazard arises when the existence of insurance encourages the insured party to take risks that increase the likelihood of an insurance payoff. For example, a person covered by burglary insurance might not take as many precautions to prevent a burglary because they know the insurance company will pay for most of the losses if a theft occurs. Adverse selection holds that the people most likely to receive large insurance payoffs are the ones who will most want to purchase insurance. For example, a person suffering from a terminal disease would want to take out the biggest life and medical insurance policies possible, thereby exposing the insurance company to potentially large losses. Both adverse selection and moral hazard can result in large losses to insurance companies because they lead to higher payouts on insurance claims. Lowering adverse selection and moral hazard to reduce these payouts is therefore an extremely important goal for insurance companies, and this goal explains the insurance practices we will discuss here.

Screening

To reduce adverse selection, insurance providers try to screen out good insurance risks from poor ones. Effective information-collection procedures are therefore an important principle of insurance management.

When you apply for auto insurance, the first thing your insurance agent does is ask you questions about your driving record (number of speeding tickets and accidents), the type of car you are insuring, and certain personal matters (age, marital status). If you are applying for life insurance, you go through a similar grilling, but you are asked even more personal questions about such things as your health, smoking habits, and your drug and alcohol use. The life insurer even orders a medical evaluation (usually done by an independent company) that involves taking blood and urine samples. Just as a bank calculates a credit score to evaluate a potential borrower, the insurer uses the information you provide to allocate you to a risk class—a statistical estimate of how likely you are to have an insurance claim. Based on this information, the insurer can decide whether to insure you or turn you down because you pose too high a risk and thus would be an unprofitable customer.

Risk-Based Premiums

Charging insurance premiums on the basis of how much risk a policyholder poses for the insurance provider is a time-honoured principle of insurance management. Adverse selection explains why this principle is so important to insurance company profitability.

To understand why an insurance provider finds it necessary to have risk-based premiums, let's look at an example of risk-based insurance premiums that at first glance seems unfair. Harry and Sally, both college students with no accidents or speeding tickets, apply for auto insurance. Normally, Harry will be charged a much higher premium than Sally. Insurance companies do this because young males have a much higher accident rate than young females do. Suppose, though, that one insurer does not base its premiums on a risk classification but rather charges a premium based on the average combined risk for males and females. Then Sally would be charged too much and Harry too little. Sally can go to another insurer and get a lower rate, while Harry will sign up for the insurance. Because Harry's premium isn't high enough to cover the accidents he is likely to have, on average the insurer will lose money on Harry. Only with a premium based on a risk classification, so that Harry is charged more, can the insurance provider make a profit and stay in business.[2]

Restrictive Provisions

Restrictive provisions in policies are an insurance management tool for reducing moral hazard. Such provisions discourage policyholders from engaging in risky activities that make an insurance claim more likely. For example, life insurers have provisions in their policies that eliminate death benefits if the insured person commits suicide within the first two years that the policy is in effect. Restrictive provisions may also require certain behaviours on the part of the insured. A company that rents motor scooters may be required to provide helmets to renters if it wishes to be covered for liabilities associated with the rental. The role of restrictive provisions is not unlike that of restrictive covenants on debt contracts as described in Chapter 8. Both serve to reduce moral hazard by ruling out undesirable behaviour.

Prevention of Fraud

Insurance providers also face moral hazard because an insured person has an incentive to lie to the insurer and seek a claim even if the claim is not valid. For example, a person who has not complied with the restrictive provisions of an insurance contract may still submit a claim. Even worse, a person may file a claim for an event that did not actually occur. Thus an important management principle for insurance providers is that they must conduct investigations to prevent fraud so that only policyholders with valid claims receive compensation.

Cancellation of Insurance

Being prepared to cancel policies is another insurance management tool. Insurers can discourage moral hazard by threatening to cancel a policy when the insured person engages in activities that make a claim more likely. If your auto insurance company makes it clear that its policy is to cancel the coverage of any driver who gets too many speeding tickets, you will be less likely to speed.

[2]Note that the example here is in fact the lemons problem described in Chapter 8.

Deductibles

The **deductible** is a fixed amount by which the insured's loss is reduced when a claim is paid out. A $500 deductible on an auto policy, for example, means that if you suffer a loss of $1200 because of an accident, the insurer will pay you only $700. Deductibles are an additional management tool that helps insurance providers reduce moral hazard. With a deductible, you experience a loss along with the insurer when you make a claim. Because you too (not just the insurance company) stand to lose if you have an accident, you have an incentive to drive more carefully. A deductible thus encourages the policyholder to act more in a way that is profitable for the insurer; moral hazard has been reduced. And because moral hazard has been reduced, the insurance provider can lower the premium by more than enough to compensate the policyholder for the existence of the deductible. The deductible also functions to eliminate the administrative costs of handling small claims by forcing the insured to bear these losses.

Coinsurance

When a policyholder shares a percentage of the losses with the insurer, the arrangement is called **coinsurance**. For example, some medical insurance plans provide coverage for 80% of medical bills, and the insured person pays 20% after a certain deductible has been met. Coinsurance works to reduce moral hazard in exactly the same way that a deductible does. A policyholder who suffers a loss along with the insurer has less incentive to take actions, such as going to the chiropractor unnecessarily, that will generate more claims. Coinsurance is thus another useful management tool for insurance providers.

Limits on the Amount of Insurance

Another important principle of insurance management is that limits should be placed on the amount of insurance provided, even if a customer is willing to pay for more coverage. The higher the insurance coverage, the more the insured person stands to gain from engaging in risky activities that make an insurance payout more likely, and hence the greater the moral hazard. For example, if Zelda's car is insured for more than its true value, she might not take proper precautions to prevent its theft, such as making sure that the key is always removed or putting in an alarm system. If her car is stolen, she will come out ahead because the excessive insurance payment will enable her to buy an even better car. By contrast, if the amount of insurance coverage is lower than the value of her car, she will suffer a loss if the car is stolen, and thus she will take precautions to prevent this from happening. Insurance providers must always make sure that the coverage is not so high that moral hazard leads to large losses.

Summary

Effective insurance management consists of several practices: information collection and screening of potential policyholders, risk-based premiums, restrictive provisions, prevention of fraud, insurance cancellation provisions, deductibles, coinsurance, and limits on the amount of insurance. All of these practices reduce moral hazard and adverse selection by making it harder for policyholders to benefit if they engage in activities that increase the amount and likelihood of claims. With less to gain, the poor insurance risks (those who are more likely to engage in these activities in the first place) see less benefit from the insurance and thus are less likely to seek it out. ◆

S2.2 PENSION FUNDS

LO S2.2 Summarize the distinctions between defined-benefit and defined-contribution pensions and the key features of private and public pension plans.

In performing the financial intermediation function of asset transformation, pension funds provide the public with another kind of protection: regular income payments during retirement. Employers, unions, and private individuals can set up pension plans, which acquire funds through contributions from the plan's participants. Pension plans have grown in importance in recent years, with their share of total financial intermediary assets rising significantly. Federal tax policy has been a major factor behind the rapid growth of pension funds because employer contributions to employee pension plans are tax-deductible. Furthermore, tax policy has also encouraged employee contributions to pension funds by making them tax-deductible and enabling self-employed individuals to set up their own tax-sheltered pension plans and individual retirement accounts.

Because the benefits paid out of a pension fund each year are highly predictable, pension funds invest in long-term securities, with the bulk of their asset holdings in bonds, stocks, and long-term mortgages. The key management issues associated with pension funds revolve around asset management. Pension fund managers try to hold assets with high expected returns and also try to reduce risk through diversification. The investment strategies of pension plans have changed radically over time. In the aftermath of World War II, most pension fund assets were held in government bonds. However, the strong performance of stocks in the 1950s and 1960s afforded pension plans higher returns, causing them to shift their portfolios into stocks, which currently account for approximately two-thirds of their assets. As a result, pension plans now have a much stronger presence in the stock market.

Although the purpose of all pension plans is the same, they can differ in a number of attributes. First is the method by which payments are made: If the benefits are determined by the contributions into the plan and their earnings, the pension is a **defined-contribution plan**; if future income payments (benefits) are set in advance by the pension plan, the pension is a **defined-benefit plan**. Defined-benefit plans fall into two categories: A defined-benefit plan is **fully funded** if contributions into the plan and their earnings over the years are sufficient to pay out the defined benefits when they come due. If the contributions and earnings are not sufficient to pay out the defined benefits, the plan is **underfunded**. For example, if Jane Brown contributes $100 per year to her pension plan and the interest rate is 10%, then after 10 years her contributions and their earnings will be worth $1753.[3] If the defined benefit on her pension plan is $1753 or less after 10 years, then the plan is fully funded, because her contributions and earnings will be sufficient to pay this amount in full. But if the defined benefit on her plan is $2000, the plan is underfunded, because her contributions and their earnings do not cover this amount.

[3]The $100 contributed in year 1 would be worth $100 $(1 + 0.10)^{10}$ = $259.37 at the end of 10 years; the $100 contributed in year 2 would be worth $100 $(1 + 0.10)^{9}$ = $235.79 at the end of 10 years; and so on, until the $100 contributed in year 10 will be worth $100 (1 + 0.10) = $110. Adding these together, we get the total value of these contributions and their earnings at the end of 10 years:

$$\$259.37 + \$235.79 + \$214.36 + \$194.87 + \$177.16 + \$161.05 + \$146.41 + \$133.10 + \$121.00 + \$110.00 = \$1753.11$$

A second characteristic of pension plans is their *vesting*, the length of time that a person must be enrolled in the pension plan (i.e., must be either a member of the union or an employee of the company) before they are entitled to receive benefits. Typically, firms require that an employee work five years for the company before they become vested and are qualified to receive pension benefits; if the employee leaves the firm before the five years are up, because the employee either quits or is fired, all rights to the benefits are lost.

Public Pension Plans

The most important public pension plan is the Old Age Security (OAS) program, which also includes the Guaranteed Income Supplement (GIS) and Spouse's Allowance (SPA) programs. The OAS currently guarantees a minimum income to all Canadians aged 65 and over and in some cases also pays income to persons before age 65. In particular, it makes monthly flat payments (out of federal government revenues) to retired or disabled workers or their surviving spouses.

Other government-administered pension plans are the Canada Pension Plan (CPP) and in Québec the Québec Pension Plan (QPP), both established in 1966 and intended for persons who hold a job. They are supported by contributions from employees and their employers and cover almost all workers in Canada (the participation of workers aged 18 and over is compulsory). The accumulated funds of the CPP are managed by an investment board called the CPP Investment Board, and those of the QPP are managed by the *Caisse de dépôt et placement du Québec*.

When the government pension plans were established, the federal government intended to operate them like private pension plans. However, unlike private pension plans, benefits are typically paid out from current contributions, not tied closely to a participant's past contributions. This "pay-as-you-go" system led to a massive underfunding. The problems of the public pension plans could become worse in the future because of the growth in the number of retired people relative to the working population. The government has been grappling with the problems of the public pension plans for years, but the prospect of a huge bulge of new retirees has resulted in calls for radical surgery (see the FYI box "Should Public Pension Plans Be Privatized?"). In 1999, for example, the CPP was given authority to sharply increase contribution rates from 5.6% in 1999 to 9.9% in 2003. It was also given the authority to invest its accumulated assets in the market in order to earn a higher return so that future increases in contribution levels will not be needed.

Private Pension Plans

Private pension plans are voluntary, employer-sponsored plans, with the contributions usually shared between employer and employee. They are registered with the federal or a provincial pension regulatory authority as well as with the Canada Revenue Agency for tax purposes. They are also known in Canada as Registered Pension Plans.

Trusteed pension plans account for the majority of registered pension plans in Canada. Typically, a pension plan sponsor (such as a government or private-sector employer) will hire a bank, a life insurance company, or a pension fund manager to manage the fund for the benefit of the plan members. The largest trusteed pension funds in Canada are the Ontario Teachers' Pension Plan Board and the Ontario Municipal Employees' Retirement Board. Among the largest pension fund managers are the *Caisse de dépôt et placement du Québec* and Royal Trust, a subsidiary of the Royal Bank.

FYI Should Public Pension Plans Be Privatized?

In recent years, public confidence in the public pension plans has reached new lows. Some surveys suggest that young people have more confidence in the existence of flying saucers than they do in the government's ability to pay them their public pension plan benefits when they become eligible for these benefits. Without some overhaul of the system, public pension plans will not be able to meet their future obligations. This problem is not restricted to Canada, however. Indeed, the United States, European countries, and Japan face an even greater problem because their populations are aging more quickly than the Canadian population. How to fix the public pension plans has become a hotly debated political issue in recent years.

Currently, the assets of the public pension plans, which reside in trust funds, are invested in government securities. Because stocks and corporate bonds have higher returns than government securities do, many proposals to save the public pension plans suggest investing part of the trust fund in corporate securities, a move that would partially privatize the system.

Proposals for privatization take three basic forms:

1. *Government investment of trust fund assets in corporate securities.* This plan has the advantage of possibly improving the trust funds' overall return while minimizing transaction costs because it exploits the economies of scale of the trust fund. Critics warn that government ownership of private assets could lead to increased government intervention in the private sector.
2. *Shift of trust fund assets to individual accounts that can be invested in private assets.* This option has the advantage of possibly increasing the return on investments and does not involve the government in the ownership of private assets. However, critics warn that it might expose individuals to greater risk and to transaction costs on individual accounts that might be very high because of the small size of many of these accounts.
3. *Individual accounts in addition to those in the trust funds.* This option has advantages and disadvantages similar to those of option 2 and may provide more funds to individuals at retirement. However, some increase in taxes would be required to fund these accounts.

Whether some privatization of the public pension plans occurs is an open question. In the short term, public pension plan reform is likely to involve an increase in taxes, a reduction in benefits, or both. For example, under the 1997 changes to the CPP Act, the percentage of liabilities of CPP that are funded increased from 8% to 20% in 2018.

Many companies' pension plans are underfunded because the companies plan to meet their pension obligations by taking funds from current earnings when the benefits come due. As long as companies have sufficient earnings, underfunding creates no problems, but if they do not, they may not be able to meet their pension obligations. Because of problems caused by corporate underfunding, mismanagement, fraudulent practices, and other abuses of private pension funds, these funds are heavily regulated. As with insurance companies, the regulatory system governing pension funds is split between OSFI and provincial superintendents of pensions. Pension funds administered for people working in businesses that are federal in scope (for example, railways, air transport, telecommunications, and banking) are the responsibility of OSFI. However, most trusteed pension funds are registered under provincial acts. These acts and regulations tend to be similar across provinces. They establish minimum standards for the reporting and disclosure of information, set rules for vesting, put limits on the degree of underfunding allowed, and place restrictions on investment practices.

Provincial and local governments and the federal government, like private employers, have also set up pension plans for their employees. These plans are almost identical in

operation to private pension plans and hold similar assets. Underfunding of the plans is also prevalent, and some investors in municipal bonds worry that it may lead to future difficulties in the ability of provincial and local governments to meet their debt obligations.

Personal Pension Plans

An alternative to both public pension plans and private pension plans is **personal pension plans**. These are the Registered Retirement Savings Plans (RRSPs) that Canadians set up with financial institutions. They provide tax-sheltered, self-financed retirement funds. Upon retirement, an RRSP must be converted into an annuity or a Registered Retirement Income Fund (RRIF) that provides taxable annuity payments. Since RRIFs are less popular than annuities, which may only be offered by insurance companies, the insurance industry has benefited from the introduction of the RRSP.

Another way to set money aside for retirement is the Tax-Free Savings Account (TFSA), introduced by the federal government in 2009. Unlike RRSP contributions, TFSA contributions are not tax-deductible. However, the amount contributed as well as the income earned in the account is tax-free when it is withdrawn.

S2.3 FINANCE COMPANIES

LO S2.3 List and describe the different types of finance companies.

Finance companies acquire funds by issuing commercial paper or stocks and bonds or borrowing from banks, and they use the proceeds to make loans (often for small amounts) that are particularly well suited to consumer and business needs. The financial intermediation process of finance companies often involves borrowing in large amounts but lending in small amounts—a process quite different from that of banking institutions, which collect deposits in small amounts and then often make large loans. Renamed in 1998, finance companies became known as *nondepository credit intermediaries* under the industry classification system.

A key feature of finance companies is that although they lend to many of the same customers that borrow from banks, they are virtually unregulated compared to chartered banks and near banks. Provinces regulate the maximum amount that finance companies can loan to individual consumers and the terms of the debt contract, but no restrictions are placed on how they pursue branching, the assets they hold, or how they raise their funds. The lack of restrictions enables finance companies in many cases to tailor their loans to meet customer needs better than banking institutions can.

Finance companies are of three types: sales, consumer, and business.

1. *Sales finance companies* are owned by a particular retailing or manufacturing company and make loans to consumers so that they can purchase items from that company. Sales finance companies compete directly with banks for consumer loans and are used by consumers because loans frequently can be obtained faster and more conveniently at the location where an item is purchased.
2. *Consumer finance companies* make loans to consumers who wish to buy particular items such as furniture or home appliances, make home improvements, or refinance small debts. Consumer finance companies are separate corporations or are owned

by banks. Typically, these companies make loans to consumers who cannot obtain credit from other sources, and they therefore charge higher interest rates.

3. *Business finance companies* provide specialized forms of credit to businesses by making loans and purchasing accounts receivable (bills owed to the firm) at a discount; this provision of credit is called *factoring*. For example, a dressmaking firm might have outstanding bills (accounts receivable) of $100·000 owed by the retail stores that have bought its dresses. If this firm needs cash to buy 100 new sewing machines, it can sell its accounts receivable for, say, $90·000 to a finance company, which is now entitled to collect the $100·000 owed to the firm. Besides factoring, business finance companies also specialize in leasing equipment (such as railroad cars, jet planes, and computers), which they purchase and then lease to businesses for a set number of years.

S2.4 SECURITIES MARKET OPERATIONS

LO S2.4 Summarize the roles of investment banks, securities brokers and dealers, and organized exchanges in the securities market.

The smooth functioning of securities markets, in which bonds and stocks are traded, involves several financial institutions, including securities brokers and dealers, investment banks, and organized exchanges. None of these institutions were included in our list of financial intermediaries in Chapter 2 because these institutions do not perform the intermediation function of acquiring funds by issuing liabilities and then using the funds to acquire financial assets. Nonetheless, they are important in the process of channelling funds from savers to spenders and can be thought of as "financial facilitators."

First, however, we must recall the distinction between primary and secondary securities markets discussed in Chapter 2. In a primary market, new issues of a security are sold to buyers by the corporation or government agency borrowing the funds. A secondary market then trades the securities, which after their initial primary market sale are considered to be "secondhand." *Investment banks* assist in the initial sale of securities in the primary market; *securities brokers* and *dealers* assist in the trading of securities in the secondary markets, some of which are organized into exchanges.

Investment Banking

When a corporation wishes to borrow (raise) funds, it normally hires the services of an investment banker to help sell its securities. (Despite its name, an investment bank is not a bank in the ordinary sense; that is, it is not engaged in financial intermediation that takes in deposits and then lends them out.) The top investment banks in Canada are subsidiaries of the Big Six Canadian banks, and they are known as RBC Capital Markets, TD Securities, Scotia Capital, CIBC World Markets, BMO Nesbitt Burns, and National Bank Financial.

Investment bankers assist in the sale of securities as follows. First, they advise the corporation on whether it should issue bonds or stock. If they suggest that the corporation issue bonds, investment bankers give advice on what the maturity and interest payments on the bonds should be. If they suggest that the corporation should sell stock, they give advice on what the price should be. This is fairly easy to do if the firm has prior issues of stock, called **seasoned issues**, currently selling in the market. However,

when a firm issues stock for the first time in an **initial public offering (IPO)**, it is more difficult to determine what the correct price should be. All the skills and expertise of the investment banking firm then must be brought to bear to determine the most appropriate price.

IPOs have become very important in the Canadian economy because they are a major source of financing for Internet companies, which became all the rage on Bay Street in the late 1990s. Not only have IPOs helped these companies acquire the capital they needed to substantially expand their operations, but they also have made the original owners of these firms very rich. Many a 20- to 30-year-old became an instant millionaire when their stake in an Internet company was given a high valuation after the initial public offering of shares in the company. However, with the bursting of the tech bubble in 2000, many of these same people lost much of their wealth when the value of their shares came tumbling down to earth.

When the corporation decides which kind of financial instrument it will issue, it offers the issue to **underwriters**—investment bankers that guarantee the corporation a price on the securities and then sell them to the public. If the issue is small, only one investment banking firm underwrites it (usually the original investment banking firm hired to provide advice on the issue). If the issue is large, several investment banking firms form a syndicate to underwrite the issue jointly, thus limiting the risk that any one investment bank must take. The underwriters sell the securities to the general public by contacting potential buyers, such as banks and insurance companies, directly.

The activities of investment bankers and the operations of primary markets are heavily regulated by the provinces and the federal government. The Ontario Securities Commission, for example, is responsible for administering the Ontario Securities Act, Canada's first provincial securities act, passed in 1945. Other provinces and territories have generally tended to follow Ontario's lead and passed securities acts regulating investment banking and the trading of securities. The government of Canada doesn't have a Securities Act, but portions of the Criminal Code of Canada specifically apply to securities trading.

Securities Brokers and Dealers

Securities brokers and dealers conduct trading in secondary markets. Brokers act as agents for investors in the purchase or sale of securities. Brokers' function is to match buyers with sellers, a service for which they are paid brokerage commissions. In contrast to brokers, dealers link buyers and sellers by standing ready to buy and sell securities at given prices. Therefore, dealers hold inventories of securities and make their living by selling these securities for a slightly higher price than they paid for them—that is, on the "spread" between the asked price and the bid price. This can be a high-risk business because dealers hold securities that can rise or fall in price; in recent years, several firms specializing in bonds have collapsed. Brokers, by contrast, are not as exposed to risk because they do not own the securities involved in their business dealings.

Brokerage firms engage in all three securities market activities, acting as brokers, dealers, and investment bankers. That is, the same investment banks that handle the sale of securities in the primary markets also are involved in the retail business of trading for clients on the stock exchanges. However, the provinces and the federal government regulate the investment banking operation of the firms and also restrict brokers and dealers from misrepresenting securities and from trading on *insider information,* nonpublic information known only to the managers of a corporation.

Organized Exchanges

As discussed in Chapter 2, secondary markets can be organized either as over-the-counter markets, in which trades are conducted using dealers, or as organized exchanges, in which trades are conducted in one central location. The largest of the organized stock exchanges in Canada is the Toronto Stock Exchange, owned and operated by the TMX Group. It was established in 1852 and today boasts the third most active stock exchange in North America, after the New York Stock Exchange and Nasdaq. The TMX Group also encompasses the TSX Venture Exchange and (since 2008) the Montréal Exchange, which provides interest-rate derivatives trading and clearing, index derivatives trading and clearing, and equity derivatives trading and clearing.

Organized stock exchanges actually function as a hybrid of an auction market (in which buyers and sellers trade with each other in a central location) and a dealer market (in which dealers make the market by buying and selling securities at given prices). Securities are traded on the floor of the exchange with the help of a special kind of dealer-broker called a *specialist*. A specialist matches "buy" and "sell" orders submitted at the same price and so performs a brokerage function. However, if buy and sell orders do not match up, the specialist buys stocks or sells from a personal inventory of securities, in this manner performing a dealer function. By assuming both functions, the specialist maintains orderly trading of the securities for which they are responsible.

The growing internationalization of capital markets has encouraged another trend in securities trading. Increasingly, Canadian companies are being listed on U.S. stock exchanges, and the markets are moving toward international trading of stocks, 24 hours a day.

S2.5 MUTUAL FUNDS

LO S2.5 Describe the role and activities of mutual funds in financial intermediation.

Mutual funds are financial intermediaries that pool the resources of many small investors by selling them shares and using the proceeds to buy securities. Through the asset transformation process of issuing shares in small denominations and buying large blocks of securities, mutual funds can take advantage of volume discounts on brokerage commissions and purchase diversified holdings (portfolios) of securities. Mutual funds allow the small investor to enjoy the benefits of lower transaction costs in purchasing securities and to take advantage of the reduction of risk by diversifying the portfolio of securities held. Many mutual funds are run by brokerage firms, but others are run by banks or independent investment advisers such as Fidelity or Investors Group.

Mutual funds are primarily held by households (nearly 90%), with the rest held by financial institutions and nonfinancial businesses. Mutual funds have become increasingly important in household savings due primarily to the booming stock market during the 1990s.

Mutual funds, pension funds, and hedge funds, known jointly as *institutional investors*, now control more than 25% of the outstanding stock in Canada and the United States. Increased ownership of stocks means that institutional investors have more clout with corporate boards, often forcing changes in leadership or in corporate

policies. Particularly controversial recently has been a type of institutional investor called a **sovereign wealth fund**. Sovereign wealth funds are state-owned investment funds that invest in foreign assets (see the FYI box "Sovereign Wealth Funds: Are They a Danger?").

Mutual funds are structured in two ways. The more common structure is an **open-end fund**, from which shares can be redeemed at any time at a price that is tied to the asset value of the fund. A mutual fund also can be structured as a **closed-end fund**, in which a fixed number of nonredeemable shares are sold at an initial offering and then are traded like a common stock. The market price of these shares fluctuates with the value of the assets held by the fund. In contrast to the open-end fund, however, the price of the shares may be above or below the value of the assets held by the fund, depending on factors such as the liquidity of the shares or the quality of the management. The greater popularity of the open-end funds is explained by the greater liquidity of their redeemable shares relative to the liquidity of the nonredeemable shares held by closed-end funds.

FYI Sovereign Wealth Funds: Are They a Danger?

Sovereign wealth funds have been around a long time: The first, the Kuwait Investment Authority, was established in 1953. When governments accumulate a substantial amount of foreign exchange earnings, as has happened in oil-rich countries, they often recognize that it would be wiser to put these earnings into investments in foreign countries rather than keep them at home. The largest sovereign wealth funds are the Abu Dhabi Investment Authority, the Government Pension Fund of Norway, the Government of Singapore Investment Corporation, the Kuwait Investment Authority, the China Investment Corporation, Singapore's Temasek Holdings, and the Stabilization Fund of the Russian Federation.

Up until recently, these funds were relatively uncontroversial because they were comparatively small and were primarily invested in government bonds issued by industrialized countries. In recent years, however, sovereign wealth funds have grown in size and, in the search for higher returns, they have invested in a much broader set of assets. This shift in size—they now hold over $7 trillion in assets—and focus has led to serious concerns about them among industrialized countries.

One concern is that as the sizes of these funds increase, they may play a more important role in asset markets. Because some of these funds are very large—Norway's, Abu Dhabi's, and China's sovereign wealth funds each have over $800 billion of assets—the decision of one fund to pull out of a particular asset market could cause market instability. Sovereign wealth funds also raise national security issues because they might use their investments for political purposes. They might buy up strategically important industries or use their clout to get political concessions. This is a particular concern because the governments of Russia, China, and Arab countries control many of the largest of these funds. A third concern is that many of these funds, with the exception of the Norwegian fund, provide very little information about their operations and the assets in which they invest.

Although sovereign wealth funds do pose some dangers, these dangers are probably overplayed. Xenophobia often plays well in politics, and foreign purchases of domestic assets is often prevented under the banner of national security in order to protect domestic companies from unwanted takeovers. The lack of transparency associated with some of these large funds, however, is a serious problem. This is why organizations like the International Monetary Fund and the Organisation for Economic Co-operation and Development have proposed the implementation of rules to increase the amount of information these funds must disclose to the markets.

Originally, shares of most open-end mutual funds were sold by salespeople (usually brokers) who were paid a commission. Because this commission is paid at the time of purchase and is immediately subtracted from the redemption value of the shares, these funds are called **load funds**. Most mutual funds are currently **no-load funds**; they are sold directly to the public with no sales commissions. In both types of funds, the managers earn their living from management fees paid by the shareholders. These fees amount to approximately 0.5% of the asset value of the fund per year.

Mutual funds are regulated by a variety of agencies. As securities distributed by investment dealers and financial advisers, they fall within provincial jurisdiction. As services provided by financial institutions, they fall under the Bank Act, Trust and Loan Companies Act, and Insurance Companies Act, and are the responsibility of OSFI. Moreover, each mutual fund is a member of a self-regulating organization such as the Investment Industry Association of Canada and the Mutual Fund Dealers Association of Canada. Regulations require periodic disclosure of information on these funds to the public and restrictions on the methods of soliciting business.

Money Market Mutual Funds

An important addition to the family of mutual funds born of the financial innovation process described in earlier chapters is the *money market mutual fund*. Recall that this type of mutual fund invests in short-term debt (money market) instruments of very high quality, such as Treasury bills, commercial paper, and bank certificates of deposit. There is some fluctuation in the market value of these securities, but because their maturity term is typically less than six months, the change in the market value is small enough that these funds allow their shares to be redeemed at a fixed value. (Changes in the market value of the securities are figured into the interest paid out by the fund.)

S2.6 HEDGE FUNDS

LO S2.6 Summarize the key distinctions between hedge funds and mutual funds.

Hedge funds are a special type of investment fund, with estimated assets of more than $3 trillion. Well-known hedge funds include Tudor Investment Corporation and the Quantum Group of funds associated with George Soros. Investors in hedge funds, who are limited partners, give money to managing (general) partners to invest on their behalf. Several features distinguish hedge funds from mutual funds. Hedge funds have a minimum investment requirement between $100 000 and $20 million, with the typical minimum investment being $1 million. In the United States, federal law mandates that hedge funds have no more than 99 investors (limited partners), and these partners must have steady annual incomes of $200 000 or more or a net worth of $1 million, excluding their homes. Hedge funds are largely unregulated by the government, on the theory that the rich can look out for themselves. Many of the 5000 hedge funds are located offshore to escape regulatory restrictions.

Hedge funds also differ from traditional mutual funds in that they usually require that investors commit their money for long periods of time, often several years. This requirement is meant to give managers breathing room to pursue long-run strategies.

Hedge funds typically charge large fees to investors. The typical fund charges a 2% annual fee on the assets it manages, plus 20% of profits.

The term *hedge fund* is highly misleading, because the word "hedge" typically implies the use of strategies to avoid risk. Despite their name, these funds can and do take big risks. Many hedge funds engage in what are called "market-neutral" strategies wherein they buy a security, such as a bond, that seems cheap and sell an equivalent amount of a similar security that appears to be overvalued. If interest rates as a whole go up or down, the fund is hedged, because the decline in value of one security is matched by the rise in value of the other. However, the fund is speculating on whether the spread between the prices on the two securities will move in the direction predicted by the fund managers. If the fund bets wrongly, it can lose a lot of money, particularly if it has highly leveraged its positions—that is, has borrowed heavily against these positions so that its equity stake is small relative to the size of its portfolio.

S2.7 PRIVATE EQUITY AND VENTURE CAPITAL FUNDS

LO S2.7 Define private equity and venture capital, and summarize their advantages as investment funds.

Another type of investment fund is the **private equity fund**, which makes long-term investments in companies that are not traded in public markets and has a similar structure to the hedge fund. In a private equity fund, investors who are limited partners (e.g., high-wealth individuals, pension funds, financial institutions, and university endowments) place their money with the managing (general) partners who make the private equity investments. Private equity funds are of two types. **Venture capital funds** make investments in new startup businesses, often in the technology industry. *Capital buyout funds*, on the other hand, make investments in established businesses. In many cases, capital buyout funds buy publicly traded firms through a so-called *leveraged buyout*, in which the publicly traded firm is taken private when the fund buys all of its shares and finances the purchase by increasing the leverage (debt) of the firm. Among the best-known companies that set up venture capital and capital buyout funds are KKR (Kohlberg, Kravis, Roberts and Co.), Bain Capital, and Blackstone Group.

Investing in private equity funds has several advantages over investing in publicly traded companies. First, private companies are not subject to controversial and costly regulations. Second, managers of private companies do not feel under pressure to produce immediate profits that managers at publicly traded companies do and thus can manage their companies with their eyes on longer-term profitability. Third, because private equity funds give their managers larger stakes in the firm than is usually the case with publicly traded corporations, these managers have greater incentives to work hard to maximize the value of the firm. Fourth, private equity overcomes the free-rider problem that we discussed in Chapter 8. In contrast to publicly traded companies, which have a diverse set of owners who are happy to free-ride off of each other, venture capital and capital buyout funds are able to garner almost all of the benefits of monitoring the firm and therefore have strong incentives to make sure the firm is run properly.

In both venture capital and capital buyout funds, once the startup or the purchased company is successful, the fund earns its returns either by selling the firm

to another company or by selling it off to the public through an IPO. The managing partners of private equity funds are well compensated for their activities: Like the partners of hedge funds, they typically earn approximately a 2% fee for management of the equity fund investments and also earn 20% of the profits, an amount that is called *carried interest*.

Both venture capital and capital buyout funds have been highly profitable. Venture capital firms have been an especially important driver of economic growth in recent years because they have funded so many successful high-tech firms, including Apple Computer, Cisco Systems, Genentech, Microsoft, and Sun Microsystems (acquired by Oracle).

S2.8 GOVERNMENT FINANCIAL INTERMEDIATION

LO S2.8 Describe the two types of government financial intermediation.

The government has become involved in financial intermediation in two basic ways: by supplying government guarantees for private loans and by setting up credit agencies that directly engage in financial intermediation.

Crown Finance Companies

To promote housing and community development, in 1946 the government created the Canada Mortgage and Housing Corporation (CMHC) to provide funds to the mortgage market by borrowing from the federal government and also from the private sector by issuing mortgage-backed securities. The CMHC is not a bank; it doesn't take deposits and is not governed by the Bank Act, but as a financial intermediary it makes direct loans and investments primarily for social housing.

Agriculture is another area in which government financial intermediation plays an important role. Farm Credit Canada (FCC), headquartered in Regina, was set up as a Crown corporation in 1959 and is the successor of the Canadian Farm Loan Board, which was established in 1927 to help Canadian farmers. It makes direct loans to new and established farmers for any agricultural or farm-related operation, including the purchase of land, equipment, and livestock. It sources its funds directly from the federal government.

To stimulate the export of Canadian goods and services, Export Development Canada (EDC) was established in 1969 as the successor to the Export Credits Insurance Corporation, which dated from 1944. The EDC is a Crown corporation wholly owned by the Canadian government. With its head office in Ottawa, the EDC provides loans to Canadian exporters to finance the working capital buildup associated with international trade. It also provides intermediate-term, low-interest-rate loans to foreign concerns for the purchase of Canadian goods, equipment, and services.

To promote and assist in the establishment and development of business enterprises in Canada, in 1995 the government created the Business Development Bank of Canada (BDC), headquartered in Montreal. It is the successor to the Federal Business Development Bank, which had been set up in 1975 to succeed the Industrial Development Bank, which dated from 1944. Prior to 2008, the BDC issued its own short- and long-term debt in domestic and foreign financial markets and used the proceeds to make loans to small and medium-sized businesses. Since 2008, the BDC has been funded directly from the federal government.

Government-Sponsored Enterprises in the United States

The U.S. government has also created a number of government agencies that provide funds, either directly or indirectly, to the mortgage market. Three agencies—the Government National Mortgage Association (GNMA, or "Ginnie Mae"), the Federal National Mortgage Association (FNMA, or "Fannie Mae"), and the Federal Home Loan Mortgage Corporation (FHLMC, or "Freddie Mac")—provide funds to the U.S. mortgage market by selling bonds and using the proceeds to buy mortgages or mortgage-backed securities. Except for Ginnie Mae, which is a federal agency and thus is an entity of the U.S. government (like CMHC, FCC, EDC, and BDC are in Canada), the other agencies, known as **government-sponsored enterprises (GSEs)**, are federally sponsored agencies that function as private corporations with close ties to the government. Although the U.S. government does not explicitly back the debt of the GSEs, as is the case for government-sponsored Treasury bonds, in practice the federal government in the United States has not allowed a default on their securities.

Unfortunately, the implicit government backing of GSE debt has led to a moral hazard problem. Because the U.S. government in effect guarantees GSE debt, market discipline to limit excessive risk taking by GSEs is quite weak. The GSEs therefore have incentives to take on excessive risk, and this is exactly what they did, leaving the taxpayer holding the bag. The recent bailout of Fannie Mae and Freddie Mac by the U.S. government involved US$200 billion of government funds (see the FYI box "The Global Financial Crisis and the Bailout of Fannie Mae and Freddie Mac").

FYI The Global Financial Crisis and the Bailout of Fannie Mae and Freddie Mac

Because it encouraged excessive risk taking, the peculiar structure of Fannie Mae and Freddie Mac—private companies sponsored by the U.S. government—was an accident waiting to happen. Earlier editions of this textbook, as well as many economists, predicted exactly what came to pass: a government bailout of both companies, with huge losses for American taxpayers.

As we learned in Chapter 9, when financial institutions are backed by a government safety net, appropriate government regulation and supervision are needed to make sure these institutions do not take on excessive risk. Legislation passed in 1992 assigned Fannie and Freddie a federal regulator and supervisor, the Office of Federal Housing Enterprise Oversight (OFHEO), but this regulator was quite weak and had only a limited ability to rein these companies in. The weakness of OFHEO was not surprising. These GSEs had strong incentives to resist effective regulation and supervision because such supervision would cut into their profits. And they did resist regulation, often flagrantly: Fannie and Freddie were legendary for their lobbying machine in the U.S. Congress, about which they were not apologetic.

In 1999, Franklin Raines, Fannie's CEO at the time, said, "We manage our political risk with the same intensity that we manage our credit and interest-rate risks."* Between 1998 and 2008, Fannie and Freddie jointly spent over US$170 million on lobbyists, and from 2000 to 2008, they and their employees contributed over US$14 million to political campaign contributions.

The lobbying efforts paid off: Attempts by both the Clinton and Bush administrations to strengthen the GSEs' regulator, OFHEO, came to naught, and remarkably things did not change even when major accounting scandals were revealed at both firms in 2003 and 2004, scandals in which the firms cooked the books to smooth out their earnings. (It was only in July 2008, after the cat was let out of the bag and Fannie and Freddie were in serious trouble, that legislation was passed to put into place a stronger regulator, the Federal Housing Finance Agency, which superseded OFHEO.)

With weak regulation and strong incentives to take on risk, Fannie and Freddie grew wildly, and by 2008 these firms had purchased or were guaranteeing over US$5 trillion of mortgages and mortgage-backed securities. The accounting scandals might even have pushed Fannie and Freddie to take on more risk. With the 1992 legislation, Fannie and Freddie had been given the mission of promoting affordable housing. What better way to do this than to purchase subprime and Alt-A mortgages and mortgage-backed securities (discussed in Chapter 11)? The accounting scandals provided even stronger motivation for such purchases because they weakened political support for Fannie and Freddie, giving them even greater incentives to please the U.S. Congress and support affordable housing through the purchase of these assets. By the time the global financial crisis hit in full force, Fannie and Freddie had over US$1 trillion of subprime and Alt-A assets on their books. Furthermore, they had extremely low ratios of capital relative to their assets. Indeed, their capital ratios were far lower than those of other financial institutions like commercial banks.

By 2008, after many subprime mortgages had gone into default, Fannie and Freddie had booked large losses. Their small capital buffer meant that they had little cushion to withstand these losses, and investors started to pull out their money. Because Fannie and Freddie played such a dominant role in mortgage markets, the U.S. government could not afford to have them go out of business, since this would have had a disastrous effect on the availability of mortgage credit, which would have led to further devastating effects on the housing market. With bankruptcy imminent, the U.S. Treasury stepped in with a pledge to provide up to US$200 billion of taxpayer money to the companies, if needed. This largess did not come for free. The federal government in effect took over these companies by putting them into conservatorship, requiring that their CEOs step down, and insisting that their regulator, the Federal Housing Finance Agency, oversee the companies' day-to-day operations. In addition, the government received about US$1 billion of senior preferred stock and the right to purchase 80% of the common stock if the companies recovered. After the bailout, the prices of both companies' common stock fell to less than 2% of their value only a year earlier.

The ultimate fate of these two companies remains unclear. The sad saga of Fannie Mae and Freddie Mac illustrates how dangerous it was for the U.S. government to set up GSEs that were exposed to the classic conflict-of-interest problem: They were expected to serve two masters. As publicly traded corporations, they were expected to maximize profits for their shareholders, but as government agencies, they were obliged to work in the interests of the public. In the end, neither the public nor the shareholders was well served.

*Quoted in Nile Stephen Campbell, "Fannie Mae Officials Try to Assuage Worried Investors," *Real Estate Finance Today*, May 10, 1999.

SUMMARY

1. Insurance providers, which are regulated by OSFI and the provinces, acquire funds by selling policies that pay out benefits to the policyholder if catastrophic events occur. Property and casualty insurance companies hold more liquid assets than life insurance companies do because of greater uncertainty regarding the benefits they will have to pay out. All insurers face moral hazard and adverse selection problems, which explain the use of insurance management tools, such as information collection and screening of potential policyholders, risk-based premiums, restrictive provisions, prevention of fraud, insurance cancellation provisions, deductibles, coinsurance, and limits on the amount of insurance.
2. Pension plans provide income payments to people when they retire after contributing to the plans for many years. Federal tax policies have encouraged the rapid growth of pension funds, and these funds now play an important role in the stock market. Many pension plans are underfunded, which means that in future years they will have to pay out benefits worth more than the value of their contributions and earnings. The problem of underfunding is especially acute for public pension plans such as the CPP and QPP.
3. Finance companies raise funds by issuing commercial paper and stocks and bonds and then using the proceeds to make loans that are particularly suited to consumer

and business needs. Virtually unregulated in comparison to chartered banks and near banks, finance companies are able to tailor their loans to customer needs.

4. Investment bankers assist in the initial sales of securities in primary markets, whereas securities brokers and dealers assist in the trading of securities in the secondary markets, some of which are organized into exchanges. The provinces and the federal government regulate the financial institutions in the securities markets and ensure that adequate information reaches prospective investors.
5. Mutual funds sell shares and use the proceeds to buy securities. Open-end funds issue shares that can be redeemed at any time at a price tied to the asset value of the firm. Closed-end funds issue nonredeemable shares, which are traded like common stock. They are less popular than open-end funds because their shares are not as liquid. Money market mutual funds hold only short-term, high-quality securities, allowing shares to be redeemed at a fixed value.
6. Hedge funds are a special type of investment fund. Investors in hedge funds, who are limited partners, give money to managing (general) partners to invest on their behalf. Hedge funds require large minimum investments and usually require that investors commit their money for long periods of time.
7. Private equity funds make long-term investments in companies that are not traded publicly. Private equity funds are of two types: venture capital funds, which make investments in startups, and capital buyout funds, which make investments in established companies, often taking publicly traded firms private.
8. To provide credit to residential housing and agriculture, stimulate the export of Canadian goods and services, and assist in the development of business enterprises, the government has created a number of Crown corporations, such as CMHC, FCC, EDC, and BDC.
9. The U.S. government has also created a number of government agencies. Particularly important are government-sponsored enterprises (GSEs), which are federally sponsored agencies that function as private corporations with close ties to the government. Because the U.S. government provides an implicit guarantee for GSE debt, market discipline to limit excessive risk taking by GSEs is weak. The resulting moral hazard problem has led to major taxpayer bailouts, especially the recent bailout of Fannie Mae and Freddie Mac, which involved US$200 billion of government funds.

MyLab Economics Make the grade with MyLab Economics: Select end-of-chapter questions can be found on MyLab Economics. You can practise them as often as you want, and most feature step-by-step guided instructions to help you find the right answer.

QUESTIONS

1. If death rates were to become less predictable than they are, how would life insurance companies change the types of assets they hold?
2. Why do property and casualty insurance companies have large holdings of liquid assets but life insurance companies do not?
3. Why are all defined-contribution pension plans fully funded?
4. How can favourable tax treatment of pension plans encourage saving?
5. "In contrast to private pension plans, government pension plans are rarely underfunded." Is this statement true, false, or uncertain? Explain your answer.
6. What explains the widespread use of deductibles in insurance policies?
7. Why might insurance companies restrict the amount of insurance a policyholder can buy?
8. Why are restrictive provisions a necessary part of insurance policies?
9. If you needed to take out a loan, why might you first go to your local bank rather than to a finance company?
10. Explain why shares in closed-end mutual funds typically sell for less than the market value of the stocks they hold.
11. Why might you buy a no-load mutual fund instead of a load fund?

12. Why can a money market mutual fund allow its shareholders to redeem shares at a fixed price but other mutual funds cannot?

13. Why might government loan guarantees be a high-cost way for the government to subsidize certain activities?

14. If you like to take risks, would you rather be a dealer, a broker, or a specialist? Why?

15. Is investment banking a good career for someone who is afraid of taking risks? Why or why not?

16. How do hedge funds differ from mutual funds?

17. How do private equity funds escape the free-rider problem?

18. What are the four advantages of private equity funds? How do they help alleviate the free-rider problem?

19. How have GSEs in the United States exposed taxpayers to large losses?

APPLIED PROBLEMS

20. Your rich uncle dies, leaving you a life insurance policy worth $100 000. The insurance company offers you the option of receiving $8225 per year for 20 years, with the first payment due today, or a lump sum. Which option should you choose?

21. Kio Outfitters estimated the following probabilities of loss from past experience:

Loss	Probability (%)
$30 000	0.25
$15 000	0.75
$10 000	1.50
$5000	2.50
$1000	5.00
$250	15.00
$0	75.00

What is the probability that Kio will experience a loss of $5000 or greater? If an insurance company offers a loss policy with a $1500 deductible, what is the most that Kio will pay?

22. Paul's car slid off an icy road, causing $2500 in damage to his car. His car insurance has a $500 deductible, after which the full loss is paid. His health insurance has a $100 deductible and covers 75% of medical costs (total). What were Paul's out-of-pocket costs from the incident?

23. An employee contributes $200 a year (at the end of the year) to their pension plan. What will be the total contributions and value of the account after five years? Assume the plan earns 15% per year over the period.

3 What Is Money?

Learning Objectives

3.1 Describe what money is.

3.2 List and summarize the functions of money.

3.3 Identify different types of payment systems.

3.4 Compare and contrast the M1, M2, and M3 money supplies.

Preview

If you lived in Canada before the creation of the central bank (the Bank of Canada) in 1935, your money might have consisted primarily of gold and silver coins and paper notes, called banknotes, issued by private banks. Today, you use as money not only coins and paper bills issued by the government but also debit cards and cheques written on accounts held at banks. Money has taken different forms at different times, but it has *always* been important to people and to the economy.

To understand the effects of money on the economy, we must understand exactly what money is. In this chapter, we develop precise definitions by exploring the functions of money, looking at why and how it promotes economic efficiency, tracing how its forms have evolved over time, and examining how money is currently measured.

3.1 MEANING OF MONEY

LO 3.1 Describe what money is.

As used in everyday conversation, the word *money* can mean many things, but to economists it has a very specific meaning. To avoid confusion, we must clarify how economists' use of the word *money* differs from conventional usage.

Economists define *money* (also referred to as the *money supply*) as anything that is generally accepted as payment for goods or services or in the repayment of debts. **Currency**, consisting of paper bills and coins, clearly fits this definition and is one type of money. When most people talk about money, they're talking about currency. If, for example, someone comes up to you and says, "Your money or your life," you should quickly hand over all of your currency rather than ask, "What exactly do you mean by 'money'?"

To define money merely as currency is much too narrow a definition for economists. Because cheques are also accepted as payment for purchases, chequing account deposits are considered money as well. An even broader definition of money is needed because other items, such as savings deposits, can, in effect, function as money if they can be quickly and easily converted into currency or chequing account deposits. As you can see, no single, precise definition of money or the money supply is possible, even for economists.

To complicate matters further, the word *money* is frequently used synonymously with *wealth*. When people say, "Joe is rich—he has an awful lot of money," they probably

mean that Joe not only has a lot of currency and a high balance in his chequing account but also has stocks, bonds, four cars, three houses, and a yacht. Thus, while "currency" is too narrow a definition of money, this other popular usage is much too broad. Economists make a distinction between money in the form of currency, demand deposits, and other items that are used to make purchases, and **wealth**, the total collection of pieces of property that serve to store value. Wealth includes not only money but also other assets such as bonds, common stock, art, land, furniture, cars, and houses.

People also use the word *money* to describe what economists call *income*, as in the sentence "Sheila would be a wonderful catch; she has a good job and earns a lot of money." **Income** is a *flow* of earnings per unit of time. Money, by contrast, is a *stock*: It is a certain amount at a given point in time. If someone tells you that they have an income of $1000, you cannot tell whether they earn a lot or a little without knowing whether this $1000 is earned per year, per month, or even per day. But if someone tells you that they have $1000 in their pocket, you know exactly how much this is.

Keep in mind that the money discussed in this book refers to anything that is generally accepted as payment for goods and services or in the repayment of debts and is distinct from income and wealth.

3.2 FUNCTIONS OF MONEY

LO 3.2 List and summarize the functions of money.

Whether money is shells or rocks or gold or paper, it has three primary functions in any economy: as a medium of exchange, as a unit of account, and as a store of value. Of the three functions, its function as a medium of exchange is what distinguishes money from other assets such as stocks, bonds, and houses.

Medium of Exchange

In almost all market transactions in our economy, money in the form of currency or cheques is a **medium of exchange**; it is used to pay for goods and services. The use of money as a medium of exchange promotes economic efficiency by minimizing the time spent in exchanging goods and services. To see why, let's look at a *barter economy*, one without money, in which goods and services are exchanged directly for other goods and services.

Take the case of Ellen the Economics Professor, who can do just one thing well: give brilliant economics lectures. In a barter economy, if Ellen wants to eat, she must find a farmer who not only produces the food she likes but also wants to learn economics. As you might expect, this search will be difficult and time-consuming, and Ellen might spend more time looking for such an economics-hungry farmer than she will teaching. It is even possible that she will have to quit lecturing and go into farming herself. Even so, she may still starve to death.

The time spent trying to exchange goods or services is called a *transaction cost*. In a barter economy, transaction costs are high because people have to satisfy a "double coincidence of wants"—they have to find someone who has a good or service they want and who also wants the good or service they have to offer.

Let's see what happens if we introduce money into Ellen the Economics Professor's world. Ellen can teach anyone who is willing to pay money to hear her lecture. She can then go to any farmer (or their representative at the supermarket) and buy the food she needs with the money she has been paid. The problem of the double coincidence of

wants is avoided, and Ellen saves a lot of time, which she may spend doing what she does best: teaching.

As this example shows, money promotes economic efficiency by eliminating much of the time spent exchanging goods and services. It also promotes efficiency by allowing people to specialize in what they do best. Money is therefore essential in an economy: It is a lubricant that allows the economy to run more smoothly by lowering transaction costs, thereby encouraging specialization and division of labour.

The need for money is so strong that almost every society beyond the most primitive invents it. For a commodity to function effectively as money, it has to meet several criteria: (1) It must be easily standardized, making it simple to ascertain its value; (2) it must be widely accepted; (3) it must be divisible, so that it is easy to "make change"; (4) it must be easy to carry; and (5) it must not deteriorate quickly. Objects that have satisfied these criteria have taken many unusual forms throughout human history, ranging from wampum (strings of beads) used by Native Americans; to tobacco and whiskey, used by the early American colonists; to cigarettes, used in prisoner-of-war camps during World War II.[1] The diverse forms of money that have been developed over the years are as much a testament to the inventiveness of the human race as are the developments of tools and language.

Unit of Account

The second role of money is to provide a **unit of account**; that is, money is used to measure value in an economy. We measure the value of goods and services in terms of money, just as we measure weight in terms of kilograms or distance in terms of kilometres. To see why this function is important, let's look again at a barter economy, in which money does not perform this function. If the economy has only three goods—say, peaches, economics lectures, and movies—then we need to know only three prices to tell us how to exchange one for another: the price of peaches in terms of economics lectures (that is, how many economics lectures you have to pay for a peach), the price of peaches in terms of movies, and the price of economics lectures in terms of movies. If there were 10 goods, we would need to know 45 prices in order to exchange one good for another; with 100 goods, we would need 4950 prices; and with 1000 goods, we would need 499 500 prices.[2]

Imagine how hard it would be in a barter economy to shop at a supermarket with 1000 different items on its shelves and be faced with deciding whether chicken or fish is a better buy if the price of a kilogram of chicken were quoted as 4 kilograms of butter and the price of a kilogram of fish as 8 kilograms of tomatoes. To make it possible to compare prices, the tag on each item would have to list up to 999 different prices, and the time spent reading them would result in very high transaction costs.

[1]An extremely entertaining article on the development of money in a prisoner-of-war camp during World War II is R. A. Radford, "The Economic Organization of a P.O.W. Camp," *Economica* 12 (November 1945): 189–201.

[2]The formula for telling us the number of prices we need when we have N goods is the same formula that tells us the number of pairs when there are N items. It is

$$\frac{N(N-1)}{2}$$

In the case of 10 goods, for example, we would need

$$\frac{10(10-1)}{2} = \frac{90}{2} = 45 \text{ prices}$$

The solution to the problem is to introduce money into the economy and have all prices quoted in terms of units of that money, enabling us to quote the price of economics lectures, peaches, and movies in terms of, say, dollars. If there were only three goods in the economy, this would not be a great advantage over the barter system, because we would still need three prices to conduct transactions. But for 10 goods we would need only 10 prices; for 100 goods, 100 prices; and so on. At the 1000-goods supermarket, now only 1000 prices need to be considered, not 499 500!

We can see that using money as a unit of account lowers transaction costs in an economy by reducing the number of prices that need to be considered. The benefits of this function of money grow as the economy becomes more complex.

Store of Value

Money also functions as a **store of value**; it is a repository of purchasing power available over time. A store of value is used to save purchasing power from the time income is received until the time it is spent. This function of money is useful because most of us do not want to spend our income immediately upon receiving it but rather prefer to wait until we have the time or the desire to shop.

Money is not unique as a store of value; any asset—whether money, stocks, bonds, land, houses, art, or jewellery—can be used to store wealth. Many such assets have advantages over money as a store of value: They often pay the owner a higher interest rate than money, experience price appreciation, and deliver services such as providing a roof over one's head. If these assets are a more desirable store of value than money, why do people hold money at all?

The answer to this question relates to the important economic concept of **liquidity**, the relative ease and speed with which an asset can be converted into a medium of exchange. Liquidity is highly desirable. Money is the most liquid asset of all because it *is* the medium of exchange; it does not have to be converted into anything else to make purchases. Other assets involve transaction costs when they are converted into money. When you sell your house, for example, you have to pay a brokerage commission (usually 4%–6% of the sales price), and if you need cash immediately to pay some pressing bills, you might have to settle for a lower price if you want to sell the house quickly. Because money is the most liquid asset, people are willing to hold it even if it is not the most attractive store of value.

How good a store of value money is depends on the price level. A doubling of all prices, for example, means that the value of money has dropped by half; conversely, a halving of all prices means that the value of money has doubled. During times of inflation, when the price level is increasing rapidly, money loses value rapidly, and people become more reluctant to hold their wealth in this form. This is especially true during periods of extreme inflation, known as **hyperinflation**, in which the inflation rate exceeds 50% per month.

Hyperinflation occurred in Germany after World War I, with inflation rates sometimes exceeding 1000% per month. By the end of the hyperinflation of 1923, the price level had risen to more than 30 billion times what it had been just two years before. The quantity of money needed to purchase even the most basic items became excessive. There are stories, for example, that near the end of the hyperinflation, a wheelbarrow of cash would be required to pay for a loaf of bread. Money was losing its value so rapidly that workers were paid and then given time off on several occasions during the day to spend their wages before the money became worthless. No one wanted to hold on to money, so the use of money to carry out transactions declined, and barter became more and more dominant. Transaction costs skyrocketed and, as we would expect, output in the economy fell sharply.

3.3 EVOLUTION OF THE PAYMENTS SYSTEM

LO 3.3 Identify different types of payment systems.

We can obtain a better picture of the functions of money and the forms it has taken over time by looking at the evolution of the **payments system**, the method of conducting transactions in the economy. The payments system has been evolving over centuries and with it the form of money. At one point, precious metals such as gold were used as the principal means of payment and were the main form of money. Later, paper assets such as cheques and currency began to be used in the payments system and viewed as money. Where the payments system is heading has an important bearing on how money will be defined in the future.

Commodity Money

To obtain perspective on where the payments system is heading, it's worth exploring how it has evolved. For any object to function as money, it must be universally acceptable; everyone must be willing to take it in payment for goods and services. An object that clearly has value to everyone is a likely candidate to serve as money, and a natural choice is a precious metal such as gold or silver. Money made up of precious metals or another valuable commodity is called **commodity money**, and from ancient times until several hundred years ago, commodity money functioned as the medium of exchange in all but the most primitive societies. The problem with a payments system based exclusively on precious metals is that such a form of money is very heavy and is hard to transport from one place to another. Imagine the holes you'd wear in your pockets if you had to buy things only with coins! Indeed, for a large purchase such as a house, you'd have to rent a truck to transport the money payment.

Fiat Money

The next development in the payments system was *paper currency* (pieces of paper that function as a medium of exchange). Initially, paper currency carried a guarantee that it was convertible into coins or into a fixed quantity of precious metal. However, currency has evolved into **fiat money**, paper currency decreed by governments as legal tender (meaning that it must be accepted as legal payment for debts) but not convertible into coins or precious metal. Paper currency has the advantage of being much lighter than coins or precious metal, but it can be accepted as a medium of exchange only if there is some trust in the authorities who issue it and if printing has reached a sufficiently advanced stage that counterfeiting is extremely difficult. Because paper currency has evolved into a legal arrangement, countries can change the currency they use at will. Indeed, this is what many European countries did when they abandoned their currencies for euro bills and coins in 2002.

Major drawbacks of paper currency and coins are that they are easily stolen and can be expensive to transport in large amounts because of their bulk. To combat this problem, another step in the evolution of the payments system occurred with the development of modern banking: the invention of *cheques*.

Cheques

A cheque is an instruction from you to your bank to transfer money from your account to someone else's account when they deposit the cheque. Cheques allow transactions to take place without the need to carry around large amounts of currency. The introduction of cheques was a major innovation that improved the efficiency of the payments

system. Frequently, payments made back and forth cancel each other; without cheques, this would involve the movement of a lot of currency. With cheques, payments that cancel each other can be settled by cancelling the cheques, and no currency need be moved. The use of cheques thus reduces the transportation costs associated with the payments system and improves economic efficiency. Another advantage of cheques is that they can be written for any amount up to the balance in the account, making transactions for large amounts much easier. Cheques are also advantageous in that loss from theft is greatly reduced and because they provide convenient receipts for purchases.

Two problems arise, however, with a payments system based on cheques. First, it takes time to get cheques from one place to another, a particularly serious problem if you are paying someone in a different location who needs to be paid quickly. In addition, if you have a chequing account, you know that it often takes several business days before a bank will allow you to make use of the funds from a cheque you have deposited. If your need for cash is urgent, this feature of paying by cheque can be frustrating. Second, the paper shuffling required to process cheques is costly.

Electronic Payment

The development of inexpensive computers and the spread of the Internet now make it cheap to pay bills electronically. In the past, you had to pay a bill by mailing a cheque, but now banks provide websites where you just log on, make a few clicks, and thereby transmit your payment electronically. Not only do you save the cost of the stamp, but paying bills becomes (almost) a pleasure, requiring little effort. Electronic payment systems provided by banks now even spare you the step of logging on to pay the bill. Instead, recurring bills can be automatically deducted from your bank account. Estimated cost savings when a bill is paid electronically rather than by a cheque exceed one dollar per transaction. Electronic payment is thus becoming more and more common in Canada.

E-Money

Electronic payments technology can substitute not only for cheques but also for cash, in the form of **electronic money** (or **e-money**)—money that exists only in electronic form. The first form of e-money was the *debit card*. Debit cards, which look like credit cards, enable consumers to purchase goods and services by electronically transferring funds directly from their bank accounts to a merchant's account. Debit cards are used in almost all of the same places that accept credit cards and are now often faster to use than cash. At most supermarkets, for example, you can tap or swipe your debit card through the card reader at the checkout station, press a button, and the amount of your purchase is deducted from your bank account. Most banks, as well as companies such as Visa and MasterCard, issue debit cards, and your ATM card typically can function as a debit card.

A more advanced form of e-money is the *stored-value card*. The simplest form of stored-value card is purchased for a pre-set dollar amount that the consumer pays up front, like a prepaid phone card. The more sophisticated stored-value card is known as a **smart card**. It contains a computer chip that allows it to be loaded with digital cash from the owner's bank account whenever needed. In Asian countries, such as Japan and Korea, cellphones now have a smart card feature that raises the expression "pay by phone" to a new level. Smart cards can be loaded from ATM machines, personal computers with a smart card reader, or specially equipped telephones.

A third form of electronic money, often referred to as **e-cash**, is used on the Internet to purchase goods or services. A consumer gets e-cash by transferring cash from

FYI Are We Headed for a Cashless Society?

Predictions of a cashless society have been around for decades, but they have not come to fruition. For example, *BusinessWeek* predicted in 1975 that electronic means of payment "would soon revolutionize the very concept of money itself," only to reverse its view several years later. Despite wider use of e-money, people still make use of a lot of cash. Why has the movement to a cashless society been so slow in coming?

Although e-money might be more convenient and efficient than a payments system based on paper, several factors work against the disappearance of the paper system. First, it is very expensive to set up the computer, card reader, and telecommunications networks necessary to make electronic money the dominant form of payment. Second, electronic means of payment raise security and privacy concerns. We often hear media reports of an unauthorized hacker who has been able to access a computer database and alter the information stored there. Because this is not an uncommon occurrence, unscrupulous persons might be able to access bank accounts in electronic payments systems and steal funds by moving them from someone else's accounts into their own. The prevention of this type of fraud is no easy task, and a whole new field of computer science has developed to cope with such security issues. A further concern is that the use of electronic means of payment leaves an electronic trail that contains a large amount of personal data on buying habits. There are worries that government agencies, employers, and marketers might be able to access these data, thereby enabling them to encroach on our privacy if they choose to do so.

In conclusion, although the use of e-money will surely increase in the future, to paraphrase Mark Twain, "the reports of cash's death are greatly exaggerated."

their bank to an account with an online payment provider such as PayPal. When they want to buy something with e-cash, they use their computer, tablet, or smartphone to click the "buy" option for a particular item from a store, whereupon the e-cash is automatically transferred from their account to the merchant's account. Given the convenience of e-money, you might think that we would move quickly to a cashless society in which all payments are made electronically. However, this hasn't happened, as discussed in the FYI box "Are We Headed for a Cashless Society?"

APPLICATION Will Bitcoin or Other Cryptocurrencies Become the Money of the Future?

Cryptocurrencies are a form of e-cash that uses cryptography to make financial transactions secure. Bitcoin is considered the first successful cryptocurrency and was created by pseudonymous software developer Satoshi Nakamoto in 2009. Cryptocurrencies, such as Bitcoin, are not controlled by a single entity like a central bank but rather are created in a decentralized fashion by users who generate new units of the cryptocurrency when they use their computing power to verify and process transactions, a process referred to as "mining." Bitcoin is currently the cryptocurrency with the largest amount outstanding, but other cryptocurrencies, such as Ethereum, Ripple, Bitcoin Cash, and EOS, are widely used.

Some tech enthusiasts have characterized Bitcoin and other cryptocurrencies as the money of the future. But do these cryptocurrencies satisfy the three key functions of money that we discussed earlier? That is, do they function as a medium of exchange, a unit of account, and a store of value?

Cryptocurrencies certainly function well as a medium of exchange. They have two features that make them attractive for conducting transactions. First, the transaction fees are substantially lower than those associated with credit cards and debit cards. Second, transactions made with cryptocurrencies can be made anonymously, which is very attractive to those who want to preserve their privacy.

However, the value of cryptocurrencies undergoes huge fluctuations. For example, the price of Bitcoin has been extremely volatile, with its volatility estimated to be over seven times that of the price of gold and over eight times that of stock market indexes such as the S&P 500. In 2011, the price of a Bitcoin fluctuated between 30 cents and $32. It then rose to a peak of $255 on April 10, 2013, only to fall back down to $55 on April 17. On November 30, 2013, the price of a Bitcoin peaked at $1125, but it fell to around $200 by June 2015 and then took off to a peak value of nearly $20 000 in December 2017. By 2020, it had fallen to a value below $7000.

The high volatility of the value of the cryptocurrencies means that they do not function well as a store of value; they are just too risky. In addition, the large changes in their prices in terms of dollars from day to day mean that they cannot function as a unit of account because the prices of goods and services in terms of the cryptocurrency would also have massive fluctuations day to day. Not surprisingly, no cryptocurrency has become a unit of account: Almost no one quotes the prices of their products in terms of bitcoins or any other cryptocurrency.

Furthermore, Bitcoin and other cryptocurrencies, which are hard to trace, are used by criminal enterprises operating on the so-called dark web, websites that are hard to find by browsing. Governments may therefore restrict the use of these cryptocurrencies in the future. China has already outlawed the use of Bitcoin as a currency.

Bitcoin and other cryptocurrencies do not satisfy two of the three key functions of money. Despite the hype, our understanding of the functions of money strongly suggests that cryptocurrencies will not be the money of the future. However, some of cryptocurrencies' technology, which enables users to conduct electronic transactions cheaply, may become a feature of future electronic payments systems. Indeed, central banks are contemplating issuing their own forms of digital currencies that will have many features of cryptocurrencies but will be fixed in value to a unit of account, such as the U.S. dollar. ◆

3.4 MEASURING MONEY

LO 3.4 Compare and contrast the M1, M2, and M3 money supplies.

The definition of money as anything that is generally accepted in payment for goods and services tells us that money is defined by people's behaviour. An asset is considered money only if people believe it will be accepted by others when making payment. As we have seen, many different assets have performed this role over the centuries, ranging from gold to paper currency to chequing accounts. For this reason, a behavioural definition does not tell us which assets in our economy should be considered money. To measure money, we need a precise definition that tells us exactly which assets should be included.

TABLE 3-1 Measures of Monetary Aggregates and Their Components

Asset	M1+	M1++	M2	M2+	M2++	M3
Currency outside banks	✓	✓	✓	✓	✓	✓
Personal deposits at chartered banks			✓	✓	✓	✓
Non-personal demand and notice deposits at chartered banks			✓	✓	✓	✓
Fixed-term deposits			✓	✓	✓	✓
Deposits at trust and mortgage loan companies (TMLs)				✓	✓	
Deposits at credit unions and *caisses populaires* (CUCPs)				✓	✓	
Life insurance company individual annuities				✓	✓	
Personal deposits at government-owned savings institutions				✓	✓	
Money market mutual funds				✓	✓	
Other retail instruments					✓	
Non-money market mutual funds					✓	
All chequable deposits at chartered banks, TMLs, and CUCPs	✓	✓				
All nonchequable deposits at chartered banks, TMLs, and CUCPs		✓				
Non-personal term deposits at chartered banks						✓
Foreign currency deposits of residents at chartered banks						✓

Note: Monetary aggregates exclude interbank deposits and include continuity adjustments. In January 2007, the monetary aggregates were redefined to reflect (i) the elimination of demand deposits and (ii) the inclusion of private-sector float.

The Bank of Canada's Monetary Aggregates

The Bank of Canada (the Bank), the central banking authority responsible for monetary policy in Canada, has conducted many studies on how to measure money. The problem of measuring money has recently become especially crucial because extensive financial innovation has produced new types of assets that might properly belong in a measure of money. Since 1980, the Bank of Canada has modified its measures of money several times and has settled on the following measures of the money supply, which are also referred to as **monetary aggregates** (also shown in Table 3-1).

FYI Where Are All the Dollars?

The almost $2000 of Canadian currency held per person in Canada and over $5000 of U.S. currency held per person in the United States are surprisingly large numbers. Currency is bulky, can be easily stolen, and pays no interest, so it doesn't make sense for most of us to keep a lot of it. Do you know anyone who carries $2000 in their pockets? We have a puzzle: Where are all these dollars, and who is holding them?

Criminals are one group that holds a lot of dollars. If you were engaged in illegal activity, you would not conduct your transactions with cheques because they are traceable and therefore a potentially powerful piece of evidence against you. That explains why gangsters and drug dealers conduct most of their transactions in cash. Some businesses like to retain a lot of cash because, if they operate as a cash business, it makes their transactions less traceable; thus they can avoid declaring income on which they would otherwise have to pay taxes.

Foreigners are the other group that routinely holds Canadian (and to a larger extent U.S.) dollars. In many countries, people do not trust their own currency because their country often experiences high inflation, which erodes the value of that currency; these people hold U.S. dollars as a hedge against this inflation risk. Lack of trust in the ruble, for example, has led Russians to hoard enormous amounts of U.S. dollars. In fact, more than half of U.S. dollars are held abroad.

The M1+ and M1++ Definitions of the Money Supply

Panel (a) of Figure 3-1 illustrates the narrow definitions of the money supply in Canada. The narrowest definition is called **M1+**. It includes

1. *Currency*, which is all the paper money and coins that are in circulation, where "in circulation" means not held by banks or the government. Surprisingly, there is close to $2000 cash in the hands of each person in Canada and more than US$5000 in the hands of each person in the United States (see the FYI box "Where Are All the Dollars?").
2. The value of all chequable deposits at chartered banks, trust and mortgage loans companies (TMLs), and credit unions and *caisses populaires* (CUCPs).

These assets are all extremely liquid because they can be turned into cash quickly at very little cost. A broader definition of the money supply is **M1++**. As panel (a) of Figure 3-1 shows, M1++ includes everything that is in M1+, plus all nonchequable deposits at chartered banks, TMLs, and CUCPs.

The M2, M2+, and M2++ Definitions of the Money Supply

As panel (b) of Figure 3-1 shows, the **M2** monetary aggregate includes currency, personal deposits at chartered banks, non-personal demand and notice deposits at chartered banks, and fixed-term deposits. The **M2+** aggregate includes everything that is in M2, plus deposits at TMLs, deposits at CUCPs, life insurance company individual annuities, personal deposits at government-owned savings institutions, and money market mutual funds. The **M2++** is the broadest definition of the money supply in Canada and includes everything that is in M2+, plus other retail debt instruments and non-money market mutual funds.

The M3 Definition of the Money Supply

Another definition of the money supply is called **M3**. As panel (c) of Figure 3-1 shows, M3 includes everything that is in M2, plus non-personal term deposits at chartered banks and foreign currency deposits of residents at chartered banks. The non-personal term deposits are held by provincial and municipal governments, corporations, and institutions, and also include deposits held by one bank with another, known as interbank deposits. The foreign currency deposits are mostly denominated in U.S. dollars.

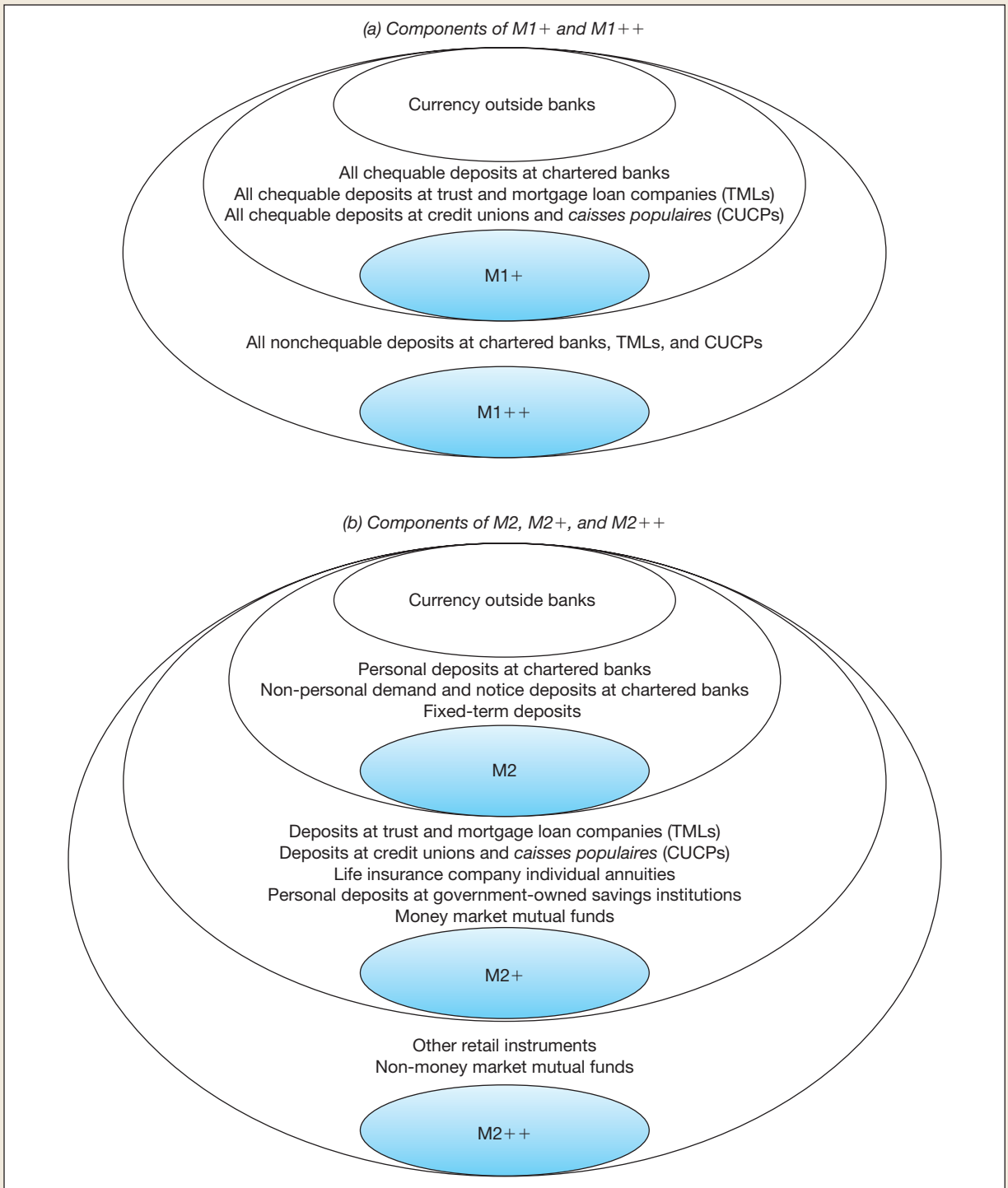

(a) Components of M1+ and M1++
Currency outside banks
All chequable deposits at chartered banks
All chequable deposits at trust and mortgage loan companies (TMLs)
All chequable deposits at credit unions and caisses populaires (CUCPs)
M1+
All nonchequable deposits at chartered banks, TMLs, and CUCPs
M1++
(b) Components of M2, M2+, and M2++
Currency outside banks
Personal deposits at chartered banks
Non-personal demand and notice deposits at chartered banks
Fixed-term deposits
M2
Deposits at trust and mortgage loan companies (TMLs)
Deposits at credit unions and caisses populaires (CUCPs)
Life insurance company individual annuities
Personal deposits at government-owned savings institutions
Money market mutual funds
M2+
Other retail instruments
Non-money market mutual funds
M2++

FIGURE 3-1 **Definitions of the Money Supply**

Continued

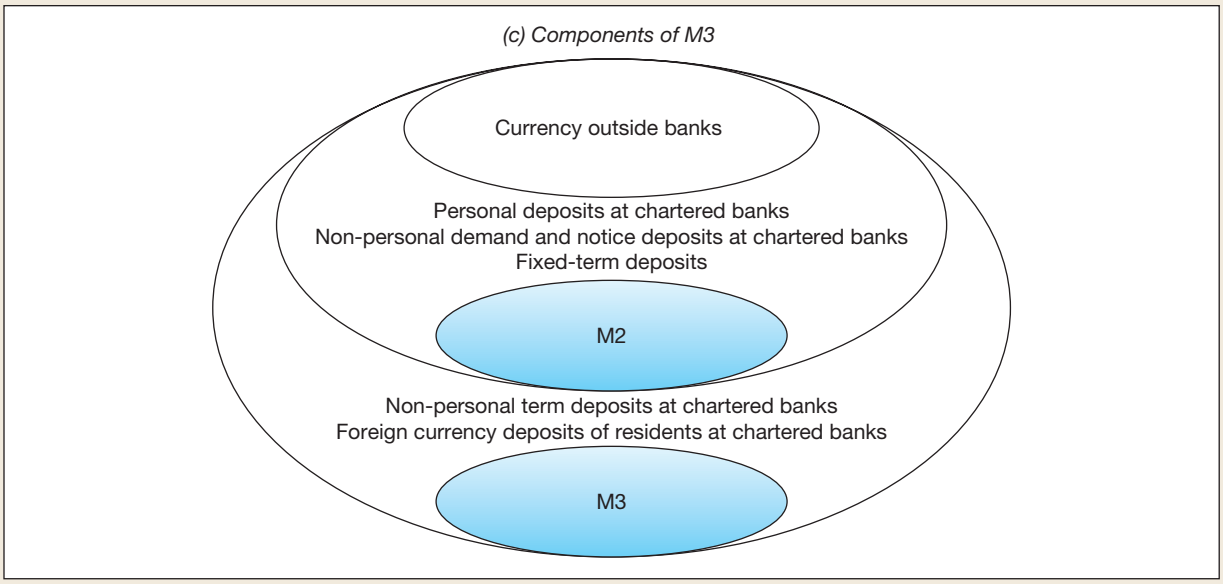

(c) Components of M3
Currency outside banks
Personal deposits at chartered banks
Non-personal demand and notice deposits at chartered banks
Fixed-term deposits
M2
Non-personal term deposits at chartered banks
Foreign currency deposits of residents at chartered banks
M3

FIGURE 3-1 (***continued***)

Notice that the M2 and M3 monetary aggregates do not include deposits with trust and mortgage loan companies and CUCPs.

Because we cannot be sure which of the monetary aggregates is the true measure of money, it is logical to wonder if their movements closely parallel one another. If they do, then using one monetary aggregate to predict future economic performance and to conduct policy will be the same as using another, and the fact that we are not sure of the appropriate definition of money for a given policy decision is not too costly. However, if the monetary aggregates do not move together, then what one monetary aggregate tells us is happening to the money supply might be quite different from what another monetary aggregate would tell us. The conflicting stories might present a confusing picture that would make it hard for policymakers to decide on the right course of action.

Figure 3-2 plots the growth rates of M2, M1++, and M2++ from 1969 to 2020. The growth rates of these three aggregates do tend to move together. Yet some glaring discrepancies exist in the movements of these aggregates. The average growth rate of money in the 1970s was 14.2% according to M2, 11.9% according to M1++, and 15% according to M2++. The average monetary growth rate in the 2000s was 7.4% according to M2, 10.2% according to M1++, and 6.6% according to M2++. Thus, the different measures of money tell a very different story about the course of monetary policy.

From the data in Figure 3-2, you can see that obtaining a single precise, correct measure of money does seem to matter and that it does make a difference which monetary aggregate policymakers and economists choose as the true measure of money.

Money as a Weighted Aggregate

The measures of the money supply listed in Table 3-1 make black-and-white decisions about whether a given asset is money by including it or excluding it. In addition, these measures are simple-sum indices in which all monetary components are assigned a constant and equal (unitary) weight. This index is M in

$$M = x_1 + x_2 + \ldots + x_n$$

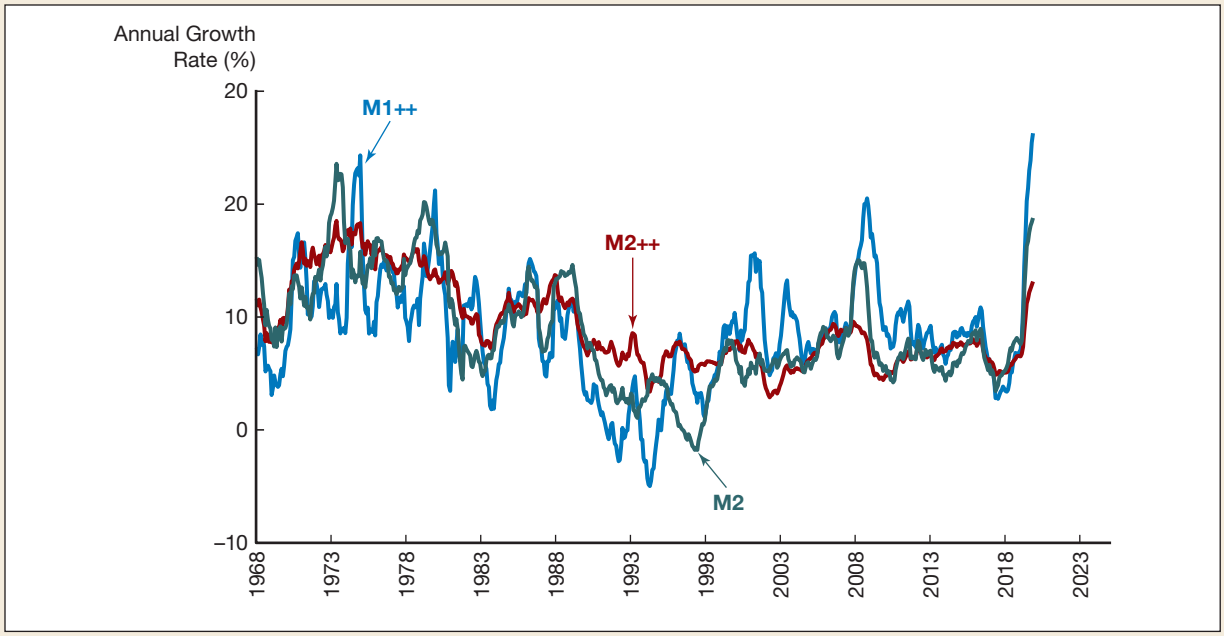


FIGURE 3-2 Growth Rates of the M2, M1++, and M2++ Aggregates, 1969–2020

The timing of the rise and fall of growth rates is roughly similar for both M1 and M2. There were periods, however, during which they moved in opposite directions, leading to conflicting recommendations about the best course of monetary policy.

Source: Statistics Canada CANSIM series V41552796, V37152, and V41552801.

where x_i is one of the n monetary components of the monetary aggregate M. This summation index implies that all monetary components contribute equally to the money total, and it views all components as dollar-for-dollar perfect substitutes. Such an index, there is no question, represents an index of the stock of nominal monetary wealth but cannot, in general, represent a valid structural economic variable for the services of the quantity of money.

Over the years there has been a steady stream of attempts at properly weighting monetary components within a simple-sum aggregate. Attention has been focused on the gains that can be achieved by a rigorous use of microeconomic theory, aggregation theory, and index number theory. This new approach to monetary aggregation led to the construction of weighted monetary aggregates.[3] These aggregates represent a viable and theoretically appropriate alternative to the simple-sum aggregates. Moreover, recent research indicates that these new measures of money seem to predict inflation and the business cycle somewhat better than more conventional measures do.[4]

[3]William Barnett, Douglas Fisher, and Apostolos Serletis, "Consumer Theory and the Demand for Money," in *Journal of Economic Literature* 30 (1992): 2086–2119, provide a state-of-the-art survey of this literature. See also William Barnett and Apostolos Serletis, *The Theory of Monetary Aggregation* (Amsterdam: North-Holland, 2000), and Apostolos Serletis, *The Demand for Money: Theoretical and Empirical Approaches,* 2nd edition (Boston: Springer, 2007).

[4]See, for example, Cosmas Dery and Apostolos Serletis, "Interest Rates, Money, and Economic Activity," in *Macroeconomic Dynamics* (2021, forthcoming).

SUMMARY

1. To economists, the word *money* has a different meaning than *income* or *wealth*. Money is anything that is generally accepted as payment for goods or services or in the repayment of debts.
2. Money serves three primary functions: as a medium of exchange, as a unit of account, and as a store of value. Money as a medium of exchange helps an economy avoid the problem of double coincidence of wants that arises in a barter economy and thus lowers transaction costs and encourages specialization and the division of labour. Money as a unit of account reduces the number of prices needed in an economy, which also reduces transaction costs. Money also functions as a store of value but performs this role poorly if it is rapidly losing value due to inflation.
3. The payments system has evolved over time. Until several hundred years ago, the payments systems in all but the most primitive societies were based primarily on precious metals. The introduction of paper currency lowered the cost of transporting money. The next major advance was the introduction of cheques, which lowered transaction costs still further. We are currently moving toward an electronic payments system in which paper is eliminated and all transactions are handled by computers. Despite the potential efficiency of such a system, obstacles are slowing the movement to a chequeless society and the development of new forms of electronic money.
4. The Bank of Canada has defined six different measures of the money supply—M1+, M1++, M2, M2+, M2++, and M3. These measures are not equivalent and do not always move together, so they cannot be used interchangeably by policymakers. Obtaining the precise, correct measure of money does seem to matter and has implications for the conduct of monetary policy.

MyLab Economics Make the grade with MyLab Economics: Select end-of-chapter questions can be found on MyLab Economics. You can practise them as often as you want, and most feature step-by-step guided instructions to help you find the right answer.

QUESTIONS

1. Why is simply counting currency an inadequate measure of money?
2. In prison, cigarettes are sometimes used among inmates as a form of payment. How is it possible for cigarettes to solve the "double coincidence of wants" problem, even if a prisoner does not smoke?
3. Three goods are produced in an economy by three individuals:

Good	Producer
Apples	Orchard owner
Bananas	Banana grower
Chocolate	Chocolatier

   If the orchard owner likes only bananas, the banana grower likes only chocolate, and the chocolatier likes only apples, will any trade between these three persons take place in a barter economy? How will introducing money into the economy benefit these three producers?
4. Why did cavemen and women not need money?
5. Most of the time it is quite difficult to separate the three functions of money. Money performs its three functions at all times, but sometimes we can stress one in particular. For each of the following situations, identify which function of money is emphasized.
   a. Brooke accepts money in exchange for performing her daily tasks at her office, since she knows she can use that money to buy goods and services.
   b. Tim wants to calculate the relative value of oranges and apples and therefore checks the price per pound of each of these goods as quoted in currency units.
   c. Maria is currently pregnant. She expects her expenditures to increase in the future and decides to increase the balance in her savings account.
6. In Brazil, a country that underwent a rapid inflation before 1994, many transactions were conducted in dollars rather than in reals, the domestic currency. Why?

7. Was money a better store of value in Canada in the 1950s than in the 1970s? Why or why not? In which period would you have been more willing to hold money?
8. Why have some economists described money during a hyperinflation as a "hot potato" that is quickly passed from one person to another?
9. Why were people in Canada in the nineteenth century sometimes willing to be paid by cheque rather than with gold, even though they knew there was a possibility that the cheque might bounce?
10. In ancient Greece, why was gold a more likely candidate for use as money than wine?
11. If you use an online payment system such as PayPal to purchase goods or services on the Internet, does this affect the M1+ money supply, the M2 money supply, both, or neither? Explain.
12. Rank the following assets from most liquid to least liquid:
    a. Chequing account deposits
    b. Houses
    c. Currency
    d. Automobiles
    e. Savings deposits
    f. Common stock
13. Which of the Bank of Canada's measures of the monetary aggregates—M2+ or M2—is composed of the most liquid assets? Which is the larger measure?
14. It is not unusual to find a business that displays a sign saying "no personal cheques, please." On the basis of this observation, comment on the relative degree of liquidity of a chequing account versus currency.
15. For each of the following assets, indicate which of the monetary aggregates (M1+, M2+, and M2++) includes them:
    a. Currency
    b. Money market mutual funds
    c. Non-money market mutual funds
    d. Chequable deposits
16. Assume that you are interested in earning some return on the idle balances you usually keep in your chequing account and decide to buy some money market mutual funds shares by writing a cheque. Comment on the effect of your action (with everything else the same) on M1+ and M2+.
17. In April 2009, the year-over-year growth rate of M1 in the United States fell to 6.1%, while the growth rate of M2 rose to 10.3%. In September 2013, the growth rate of the M1 money supply was 6.5%, while the growth rate of the M2 money supply was about 8.3%. How should Bank of Canada's policymakers interpret these changes in the growth rates of M1 and M2?
18. Suppose a researcher discovers that a measure of the total amount of debt in the Canadian economy over the past 20 years was a better predictor of inflation and the business cycle than M1+ or M2+. Does this discovery mean that we should define money as equal to the total amount of debt in the economy?

APPLIED PROBLEMS

19. Look up the M1+, M1++, M2, M2+, M2++, and M3 numbers in the Bank of Canada's Banking and Financial Statistics for the most recent one-year period. Have their growth rates been similar? What implications do their growth rates have for the conduct of monetary policy?
20. Which of the Bank of Canada's measures of the monetary aggregates, M1+, M1++, M2, M2+, M2++, or M3, is composed of the most liquid assets? Which is the largest measure?

DATA ANALYSIS PROBLEMS

1. Go to the CANSIM database and download monthly data from January 1970 to November 2020 on the M2 (gross) monetary aggregate (series V41552796).
    a. What is the average monthly monetary growth rate over the entire period?
    b. Calculate the average monthly monetary growth rate and its standard deviation over each of the last four decades.
    c. Which decade had the lowest money growth rate? The highest?
    d. Which decade had the lowest money growth volatility? The highest?

2. Go to the CANSIM database and find data on M1++ for Canada (series V37152). Calculate the annual growth rate, as $1200*\ln(M_t/M_{t-1})$, from January 2000 to November 2020. What is the average annual monetary growth rate? Its standard deviation?
3. **Real-Time Data Analysis** Go to the St. Louis Federal Reserve FRED database and find U.S. data on currency (CURRSL), demand deposits (DEMDEPSL), and other chequable deposits (OCDSL). Calculate the M1 money supply and the percentage change in M1 and in each of the three components of M1 from the most recent month of data available to the same time one year prior. Which component has the highest growth rate? The lowest growth rate?

SUPPLEMENTAL CHAPTER

3 The *ISLM* Model

Preview

In the media, you often see forecasts of GDP and interest rates by economists and government agencies. At times, these forecasts seem to come from a crystal ball, but economists actually make their predictions using a variety of economic models. One model widely used by economic forecasters is the *ISLM* model, which was developed by Sir John Hicks in 1937 and is based on the analysis in John Maynard Keynes's influential book *The General Theory of Employment, Interest, and Money*, published in 1936. The *ISLM* model explains how interest rates and total output produced in the economy (aggregate output or, equivalently, aggregate income) are determined, given a fixed price level (a reasonable assumption in the short run).

The *ISLM* model is valuable not only because it can be used in economic forecasting but also because it provides a deeper understanding of how government policy can affect aggregate economic activity. We use it to evaluate the effects of monetary and fiscal policy on the economy and to learn some lessons about how monetary policy might best be conducted.

S3.1 KEYNES'S FIXED PRICE LEVEL ASSUMPTION AND THE *IS* CURVE

In Chapter 21, we developed the *IS* curve, which is the first building block of the *ISLM* model. However, it is important to note that Keynes's analysis was fleshed out during the Great Depression period, when inflation was not a serious problem. As a result, Keynes assumed that output could change without causing a change in prices. ***ISLM analysis assumes that the price level is fixed and therefore inflation is zero***. Thus, the analysis also assumes that the nominal interest rate and the real interest rate are equal, and so there is no need to distinguish between the two. Hence in *ISLM* analysis, the *IS* curve we discussed in Chapter 21 is viewed as describing the relationship between equilibrium output and the nominal interest rate.

The *LM* Curve

We now include money and interest rates in the Keynesian framework to develop the more intricate *ISLM* model of how aggregate output is determined, in which monetary policy plays an important role. The *ISLM* model will help us understand how monetary

policy affects economic activity and interacts with fiscal policy (changes in government spending and taxes) to produce a certain level of aggregate output, and how the level of interest rates is affected by changes in investment spending as well as by changes in monetary and fiscal policy.

Equilibrium in the Market for Money: The *LM* Curve

Just as the *IS* curve is derived from the equilibrium condition in the goods market (aggregate output equals aggregate demand), the ***LM* curve** is derived from the equilibrium condition in the market for money, which occurs when the quantity of money demanded equals the quantity of money supplied. The main building block in Keynes's analysis of the market for money is the demand for money that he called *liquidity preference*. Let's briefly review his theory of the demand for money (discussed at length in Chapters 5 and 20).

Keynes's liquidity preference theory states that the demand for money in real terms M^d/P depends on income Y (aggregate output) and (nominal) interest rates i. The demand for money is positively related to income for two reasons. First, a rise in income raises the level of transactions in the economy, which in turn raises the demand for money because money is used to carry out these transactions. Second, a rise in income increases the demand for money because it increases the wealth of individuals who want to hold more assets, one of which is money. The opportunity cost of holding money is the interest sacrificed by not holding other assets (such as bonds) instead. As interest rates rise, the opportunity cost of holding money rises, and the demand for money falls. According to the liquidity preference theory, the demand for money is positively related to aggregate output and negatively related to interest rates.

Deriving the *LM* Curve

In Keynes's analysis, the level of interest rates is determined by equilibrium in the market for money (the point at which the quantity of money demanded equals the quantity of money supplied). Figure S3-1 depicts what happens to equilibrium in the market for money as the level of output changes. Because the *LM* curve is derived with the real money supply held at a fixed level, it is fixed at the level of $\overline{M}/P$ in panel (a).[1] Each level of aggregate output has its own money demand curve because as aggregate output changes, the level of transactions and wealth in the economy changes, which in turn changes the demand for money.

When aggregate output is at Y_1, the money demand curve is at $M^d(Y_1)$: The curve slopes downward because a lower interest rate means that the opportunity cost of holding money is lower, and so the quantity of money demanded is higher. Equilibrium in the market for money occurs at point 1, at which the interest rate is i_1. When aggregate output increases to Y_2, the money demand curve shifts rightward to $M^d(Y_2)$ because the higher level of output means that the quantity of money demanded is higher at any given interest rate. Equilibrium in the market for money now occurs at point 2, where the interest rate is at the higher level of i_2. Similarly, a still higher level of aggregate output Y_3 results in an even higher level of the equilibrium interest rate i_3.

Panel (b) plots the equilibrium interest rates that correspond to the different output levels, with points 1, 2, and 3 corresponding to the equilibrium points 1, 2, and 3 in panel (a). The line connecting these points is the *LM* curve, which shows the

[1]As pointed out in earlier chapters on the money supply process, the money supply is positively related to interest rates, so the M^s curve in panel (a) should actually have a positive slope. The M^s curve is drawn as a vertical line in panel (a) to simplify the graph, but allowing for a positive slope leads to identical results.

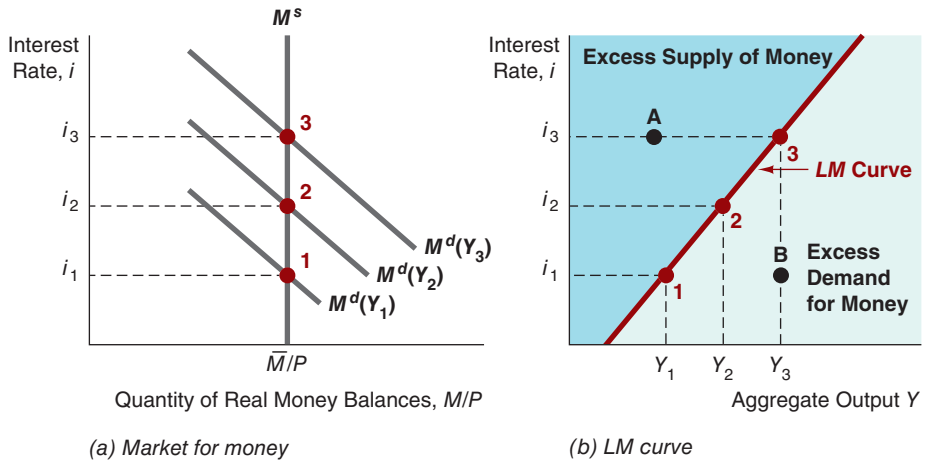


FIGURE S3-1 Deriving the *LM* Curve

Panel (a) shows the equilibrium levels of the interest rate in the market for money when aggregate output is at Y_1, Y_2, and Y_3. Panel (b) plots the three levels of the equilibrium interest rate i_1, i_2, and i_3 corresponding to these three levels of output; the line that connects these points is the *LM* curve.

combinations of interest rates and output for which the market for money is in equilibrium. The slope of the *LM* curve is positive because higher output raises the demand for money and thus raises the equilibrium interest rate.

What the *LM* Curve Tells Us The *LM* curve traces out the points that satisfy the equilibrium condition, which occurs when the quantity of money demanded equals the quantity of money supplied. For each given level of aggregate output, the *LM* curve tells us what the interest rate must be for equilibrium to exist in the market for money. As aggregate output rises, the demand for money increases and the interest rate rises so that money demanded equals money supplied and the market for money is in equilibrium.

Just as the economy tends to move toward the equilibrium points represented by the *IS* curve, it also moves toward the equilibrium points on the *LM* curve. If the economy is located in the area to the left of the *LM* curve, an excess supply of money is present. At point A, for example, the interest rate is at i_3 and aggregate output is at Y_1. The interest rate is above the equilibrium level, and people are holding more money than they wish to. To eliminate their excess money balances, they will purchase bonds, which will cause the price of the bonds to rise and their interest rate to fall. (The inverse relationship between the price of a bond and its interest rate is discussed in Chapter 4.) As long as an excess supply of money exists, the interest rate will fall until it comes to rest on the *LM* curve.

If the economy is located in the area to the right of the *LM* curve, an excess demand for money is present. At point B, for example, the interest rate i_1 is below the equilibrium level, and people want to hold more money than they currently do. To acquire this money, they will sell bonds and drive down bond prices, and the interest rate will rise. This process will stop only when the interest rate rises to an equilibrium point on the *LM* curve.

S3.2 *ISLM* APPROACH TO AGGREGATE OUTPUT AND INTEREST RATES

Now that we have derived the *IS* and *LM* curves, we can put them into the same diagram (Figure S3-2) to produce a model that enables us to determine both aggregate output and the interest rate. The only point at which the goods market and the market for money are in simultaneous equilibrium is at the intersection of the *IS* and *LM* curves, point E. At this point, aggregate output equals aggregate demand (*IS*), and the quantity of money demanded equals the quantity of money supplied (*LM*). At any other point in the diagram, at least one of these equilibrium conditions is not satisfied, and market forces move the economy toward the general equilibrium, point E.

To learn how this model works, let's consider what happens if the economy is at point A, which is on the *IS* curve but not the *LM* curve. Even though at point A the goods market is in equilibrium, so that aggregate output equals aggregate demand, the interest rate is above its equilibrium level, and so the demand for money is less than the supply. Because people have more money than they wish to hold, they will try to get rid of it by buying bonds. The resulting rise in bond prices will cause a fall in interest rates, which in turn will cause both planned investment spending and net exports to rise, thereby increasing aggregate output. The economy then will move down along the *IS* curve, and the process will continue until the interest rate has fallen to i^* and aggregate output has risen to Y^*—that is, until the economy is at equilibrium point E.

If the economy is on the *LM* curve but off the *IS* curve at point B, it will also head toward the equilibrium at point E. At point B, even though money demand

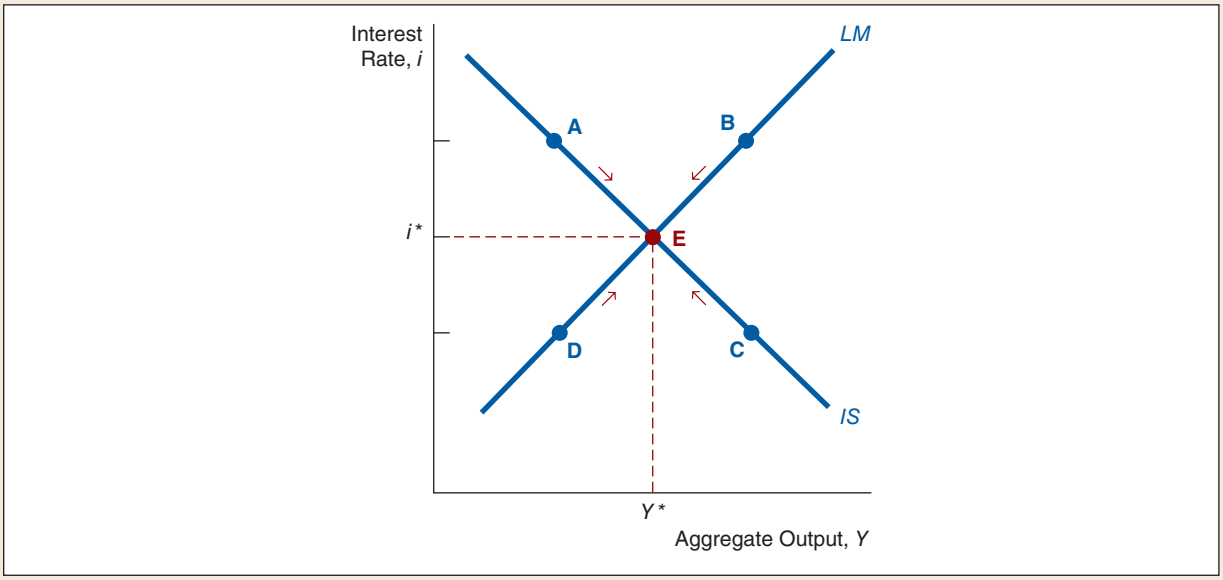


FIGURE S3-2 ***ISLM* Diagram: Simultaneous Determination of Output and the Interest Rate**

Only at point E, where the interest rate is at i^* and output is at Y^*, is there simultaneous equilibrium in both the goods market (as measured by the *IS* curve) and the market for money (as measured by the *LM* curve). At other points, such as A, B, C, or D, one of the two markets is not in equilibrium, and there will be a tendency for that market to move toward equilibrium, point E.

equals money supply, output is higher than the equilibrium level and exceeds aggregate demand. Firms are unable to sell all their output, and unplanned inventory accumulates, prompting firms to cut production and lower output. The decline in output causes the demand for money to fall, lowering interest rates. The economy then moves down along the *LM* curve until it reaches equilibrium point E.

We have finally developed a model, the *ISLM* model, that tells us how both interest rates and aggregate output are determined when the price level is fixed. Although we have demonstrated that the economy will head toward an aggregate output level of Y^*, there is no reason to assume that the economy is at full employment at this level of aggregate output. If the unemployment rate is too high, government policymakers might want to increase aggregate output to decrease unemployment. The *ISLM* apparatus indicates that they can do this by manipulating monetary and fiscal policy.

S3.3 FACTORS THAT CAUSE THE *LM* CURVE TO SHIFT

In Chapter 21, we examined the factors that cause the IS curve to shift. What factors cause the *LM* curve to shift? Only two factors cause it to shift: autonomous changes in money demand and changes in the money supply. How do changes in these two factors affect the *LM* curve?

1. *Changes in the Money Supply*. A rise in the money supply shifts the *LM* curve to the right, as shown in Figure S3-3. To see how this shift occurs, suppose the *LM* curve is initially at LM_1 in panel (a) and the Bank of Canada conducts open market

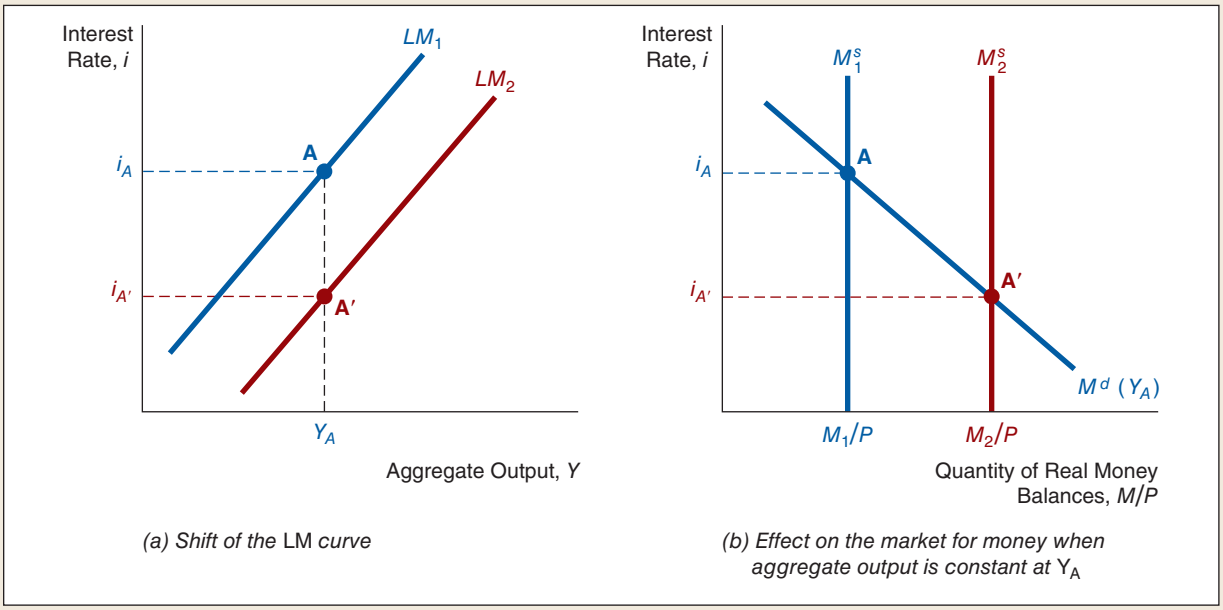


FIGURE S3-3 Shift in the *LM* Curve from an Increase in the Money Supply

The *LM* curve shifts to the right from LM_1 to LM_2 when the money supply increases because, as indicated in panel (b), at any given level of aggregate output (say, Y_A), the equilibrium interest rate falls (from point A to A′).

purchases that increase the money supply. If we consider point A, which is on the initial LM_1 curve, we can examine what happens to the equilibrium level of the interest rate, holding output constant at Y_A.

In panel (b), which depicts a supply and demand diagram for the market for money, the equilibrium interest rate initially lies at i_A, at the intersection of the supply curve for money M_1^s and the demand curve for money M^d. The rise in the quantity of money supplied shifts the supply curve rightward to M_2^s and, with output held constant at Y_A, the equilibrium interest rate falls to $i_{A'}$. In panel (a), this decline in the equilibrium interest rate from i_A to $i_{A'}$ is shown as a movement from point A to point A′. The same analysis can be applied to every point on the initial LM_1 curve, leading to the conclusion that at any given level of aggregate output, the equilibrium interest rate falls when the money supply increases. Thus LM_2 lies below and to the right of LM_1.

Reversing this reasoning, a decrease in the money supply shifts the *LM* curve to the left. A decline in the money supply results in a shortage of money at points on the initial *LM* curve. This condition of excess demand for money can be eliminated by a rise in the interest rate, which reduces the quantity of money demanded until it again equals the quantity of money supplied.

2. *Autonomous Changes in Money Demand.* The theory of portfolio choice outlined in Chapter 5 indicates that an autonomous rise in money demand (that is, a change not caused by a change in the price level, aggregate output, or the interest rate) is possible. For example, an increase in the volatility of bond returns would make bonds riskier relative to money and would increase the quantity of money demanded at any given interest rate, price level, or amount of aggregate output. The resulting autonomous increase in the demand for money would shift the *LM* curve to the

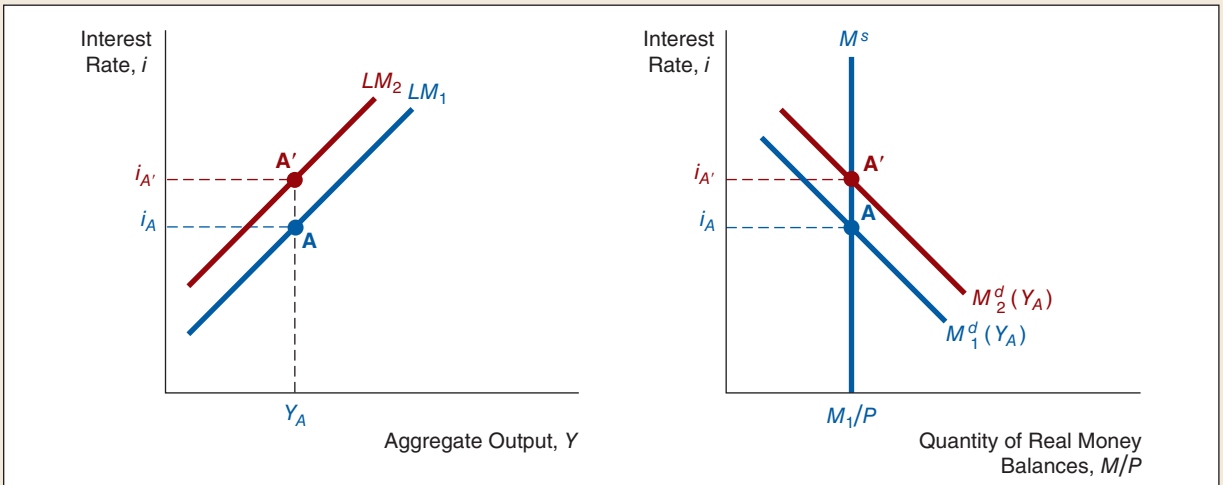


FIGURE S3-4 Shift in the *LM* Curve When Money Demand Increases

The *LM* curve shifts to the left from LM_1 to LM_2 when money demand increases because, as indicated in panel (b), the equilibrium interest rate rises (from point *A* to A′) at any given level of aggregate output (say, Y_A).

left, as shown in Figure S3-4. Consider point A on the initial LM_1 curve. Suppose a massive financial panic occurs, sending many companies into bankruptcy. Because bonds have now become a riskier asset, people will want to shift from holding bonds to holding money; they will hold more money at all interest rates and output levels. The resulting increase in money demand at an output level of Y_A is shown by the shift of the money demand curve from M_1^d to M_2^d in panel (b). The new equilibrium in the market for money indicates that if aggregate output is constant at Y_A, then the equilibrium interest rate will rise to $i_{A'}$ and the point of equilibrium will move from A to A'.

Conversely, an autonomous decline in money demand leads to a rightward shift of the *LM* curve. The fall in money demand creates an excess supply of money, which is eliminated by a rise in the quantity of money demanded resulting from a decline in the interest rate.

S3.4 CHANGES IN EQUILIBRIUM LEVEL OF THE INTEREST RATE AND AGGREGATE OUTPUT

We can now use our knowledge of the factors that cause the *IS* and *LM* curves to shift to analyze how the equilibrium levels of the interest rate and aggregate output change in response to alterations in monetary and fiscal policies.

Response to a Change in Monetary Policy

Figure S3-5 illustrates the response of output and the interest rate to an increase in the money supply. Initially, the economy is in equilibrium for both the goods market and the market for money, as indicated by point 1 at the intersection of IS_1 and LM_1.

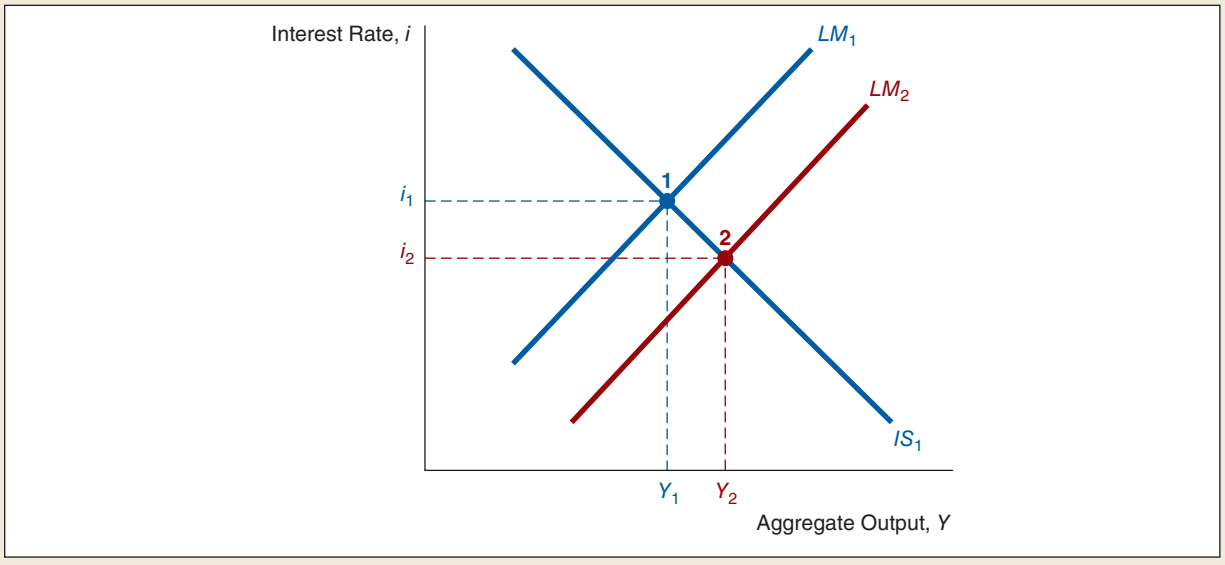


FIGURE S3-5 Response of Aggregate Output and the Interest Rate to an Increase in the Money Supply
The increase in the money supply shifts the *LM* curve to the right, from LM_1 to LM_2; the economy moves to point 2, where output has increased to Y_2 and the interest rate has declined to i_2.

Suppose that at the resulting level of aggregate output Y_1, the economy is suffering from an unemployment rate of 10%, and the Bank of Canada decides it should try to raise output and reduce unemployment by raising the money supply. Will the Bank of Canada's change in monetary policy have the intended effect?

The increase in the money supply causes the *LM* curve to shift rightward to LM_2, and the equilibrium point for both the goods market and the market for money moves to point 2 (intersection of IS_1 and LM_2). The increase in the money supply causes the interest rate to decline to i_2, as we found in Figure S3-3, and aggregate output rises to Y_2; the Bank's policy has been successful in improving the health of the economy.

To understand clearly why aggregate output rises and the interest rate declines, think about exactly what happens when we move from point 1 to point 2 in Figure S3-5. When the economy is at point 1, the increase in the money supply (rightward shift of the *LM* curve) creates an excess supply of money, resulting in a decline in the interest rate. The decline causes investment spending and net exports to increase, which in turn raises aggregate demand and aggregate output. The excess supply of money is eliminated when the economy reaches point 2 because the rise in output and the fall in the interest rate have raised the quantity of money demanded to the new, higher level of the money supply.

A decline in the money supply reverses the process; it shifts the *LM* curve to the left, causing the interest rate to rise and output to fall. Accordingly, ***aggregate output is positively related to the money supply***; aggregate output expands when the money supply increases and falls when it decreases.

Response to a Change in Fiscal Policy

Suppose that the Bank of Canada is not willing to increase the money supply when the economy is suffering from a 10% unemployment rate (at point 1). Can the federal government come to the rescue and manipulate government spending and taxes to raise aggregate output and reduce the massive unemployment?

The *ISLM* model demonstrates that it can. Figure S3-6 depicts the response of output and the interest rate to an expansionary fiscal policy (an increase in government spending or a decrease in taxes). An increase in government spending or a decrease in taxes will cause the *IS* curve to shift to IS_2, and the equilibrium point for both the goods market and the market for money will move to point 2 (intersection of IS_2 and LM_1). The result of the change in fiscal policy is a rise in aggregate output to Y_2 and a rise in the interest rate to i_2. Note the difference between an expansionary fiscal policy and an expansionary monetary policy in their effects on the interest rate. In the case of an expansionary fiscal policy, the interest rate rises, whereas in the case of an expansionary monetary policy, the interest rate falls.

Why does an increase in government spending or a decrease in taxes move the economy from point 1 to point 2, causing a rise in both aggregate output and the interest rate? An increase in government spending raises aggregate demand directly; a decrease in taxes makes more income available for spending and thus raises aggregate demand by raising consumer expenditure. The resulting increase in aggregate demand causes aggregate output to rise. The higher level of aggregate output raises the quantity of money demanded, creating an excess demand for money, which in turn causes the interest rate to rise. At point 2, the excess demand for money created by the rise in aggregate output has been eliminated by a rise in the interest rate, which lowers the quantity of money demanded.

A contractionary fiscal policy (a decrease in government spending or an increase in taxes) reverses the process described in Figure S3-5; it causes aggregate demand to fall, which shifts the *IS* curve to the left and causes both aggregate output and the interest rate to fall. ***Aggregate output and the interest rate are positively related to government spending and negatively related to taxes.***

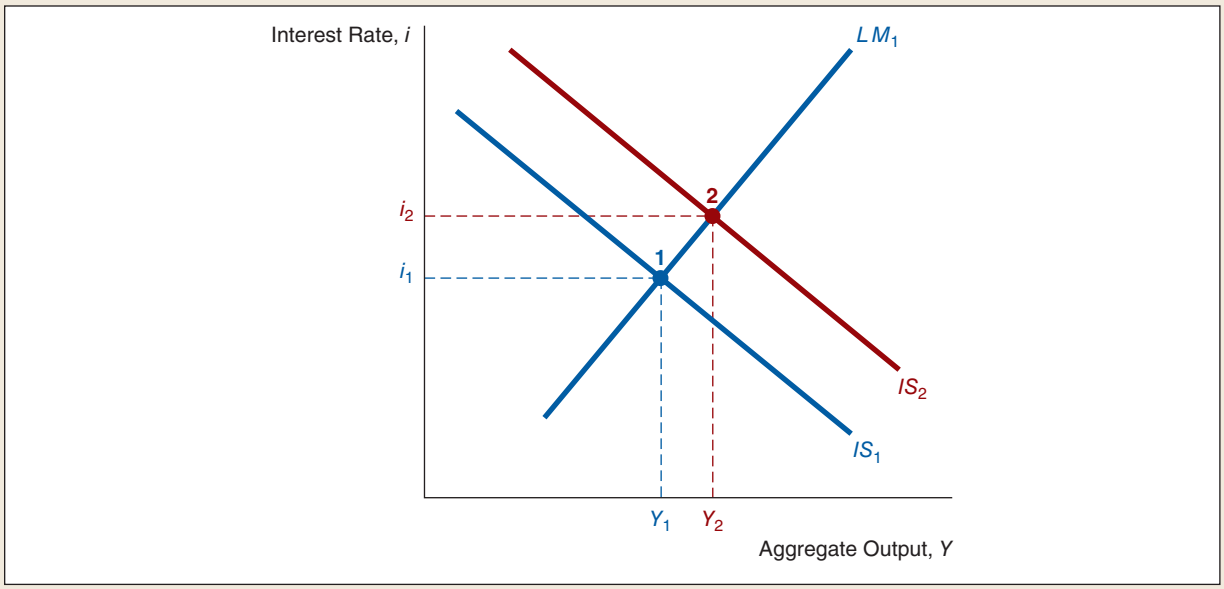


FIGURE S3-6 Response of Aggregate Output and the Interest Rate to an Expansionary Fiscal Policy
Expansionary fiscal policy (a rise in government spending or a decrease in taxes) shifts the *IS* curve to the right, from IS_1 to IS_2; the economy moves to point 2, aggregate output increases to Y_2, and the interest rate rises to i_2.

As a study aid, Summary Table S3-1 indicates the effects on aggregate output and interest rates of a change in the seven factors that shift the *IS* and *LM* curves.

APPLICATION The Covid-19 Crisis in Canada

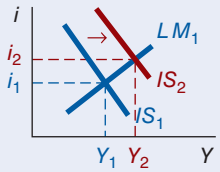
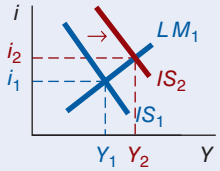
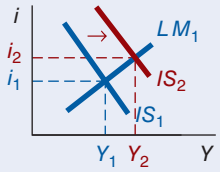
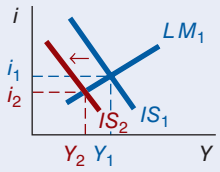
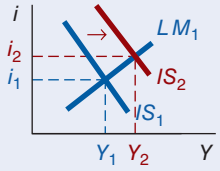
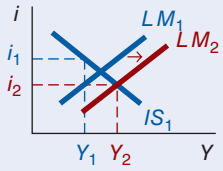
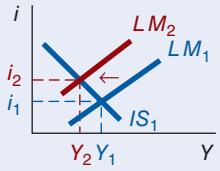
The Canadian economy was hit by a large adverse shock in March 2020, when the coronavirus pandemic led to massive lockdowns. Household and business spending collapsed, and to counter the contractionary effects on the economy of the coronavirus crisis, the federal government put in place the largest economic aid package ever, running a deficit of about $350 billion.

We can use our *ISLM* analysis to model the likely impact of this fiscal policy stimulus. Figure S3-6 shows the response of the economy to expansionary fiscal policy. The *IS* curve would shift to the right, which would raise both interest rates and aggregate output. In addition, the tax incentives for businesses in the stimulus package would lead to a rise in investment, which would provide another reason for the *IS* curve to shift to the right, and interest rates and output would rise.

Although economists have estimated that the stimulus package did increase spending a little bit and so had some of the effects described in Figure S3-6, the continuing deterioration in economic conditions overwhelmed the expansionary effects of the stimulus package. As a result, aggregate spending actually fell rather than rose. Consequently, the *IS* curve shifted to the left rather than to the right, and aggregate output and interest rates ended up falling rather than rising during this period. ◆

SUMMARY TABLE S3-1

Effects from Factors That Shift the *IS* and *LM* Curves

Factor	Autonomous Change in Factors	Response	Reason	
Consumer expenditure, C	↑	$Y\uparrow, i\uparrow$	$C\uparrow \Rightarrow Y^{ad}\uparrow \Rightarrow$ *IS* shifts right	
Investment, I	↑	$Y\uparrow, i\uparrow$	$I\uparrow \Rightarrow Y^{ad}\uparrow \Rightarrow$ *IS* shifts right	
Government spending, G	↑	$Y\uparrow, i\uparrow$	$G\uparrow \Rightarrow Y^{ad}\uparrow \Rightarrow$ *IS* shifts right	
Taxes, T	↑	$Y\downarrow, i\downarrow$	$T\uparrow \Rightarrow C\downarrow \Rightarrow Y^{ad}\downarrow \Rightarrow$ *IS* shifts left	
Net exports, NX	↑	$Y\uparrow, i\uparrow$	$NX\uparrow \Rightarrow Y^{ad}\uparrow \Rightarrow$ *IS* shifts right	
Money supply, M^s	↑	$Y\uparrow, i\downarrow$	$M^s\uparrow \Rightarrow i\downarrow \Rightarrow$ *LM* shifts right	
Money demand, M^d	↑	$Y\downarrow, i\uparrow$	$M^d\uparrow \Rightarrow i\uparrow \Rightarrow$ *LM* shifts left	

Note: Only increases (↑) in the factors are shown. The effect of decreases in the factors would be the opposite of those indicated in the "Response" column.

S3.5 EFFECTIVENESS OF MONETARY VERSUS FISCAL POLICY

Our discussion of the effects of fiscal and monetary policy suggests that a government can easily lift an economy out of a recession by implementing any of a number of policies (changing the money supply, government spending, or taxes). But how can policymakers decide which of these policies to use when faced with too much unemployment? Should they decrease taxes, increase government spending, raise the money supply, or do all three? And if they decide to increase the money supply, then by how much should they raise it? Economists do not pretend to have all the answers, and although the *ISLM* model will not clear the path to aggregate economic bliss, it can help policymakers decide which policies are likely to be most effective under certain circumstances.

Monetary Policy Versus Fiscal Policy: The Case of Complete Crowding Out

The *ISLM* model developed so far in this chapter shows that both monetary policy and fiscal policy affect the level of aggregate output. To understand when monetary policy is more effective than fiscal policy, we will examine a special case of the *ISLM* model in which money demand is unaffected by the interest rate (money demand is said to be "interest-inelastic") so that monetary policy affects output but fiscal policy does not.

Consider the slope of the *LM* curve when the demand for money is unaffected by changes in the interest rate. If point 1 in panel (a) of Figure S3-7 represents a point at which the quantity of money demanded is equal to the quantity of money supplied, then it is on the *LM* curve. If the interest rate rises to, say, i_2, the quantity of money demanded is unaffected, and it will continue to equal the *unchanged* quantity of money supplied only if aggregate output remains *unchanged* at Y_1 (point 2). Equilibrium in the market for money will occur at the same level of aggregate output regardless of the interest rate, and the *LM* curve will be vertical, as shown in both panels of Figure S3-7.

Suppose the economy is suffering from a high rate of unemployment, which policymakers would like to eliminate through either expansionary fiscal or expansionary monetary policy. Panel (a) depicts what happens when an expansionary fiscal policy (an increase in government spending or a cut in taxes) is implemented, shifting the *IS* curve to the right from IS_1 to IS_2. As you can see in panel (a), the fiscal expansion has no effect on output; aggregate output remains at Y_1 when the economy moves from point 1 to point 2.

In our earlier analysis, expansionary fiscal policy always increased aggregate demand and raised the level of output. Why does panel (a) show a different result? The answer is that because the *LM* curve is vertical, the rightward shift of the *IS* curve raises the interest rate to i_2, which causes investment spending and net exports to fall enough to offset completely the increased spending resulting from the expansionary fiscal policy. Put another way, increased spending resulting from expansionary fiscal policy has *crowded out* investment spending and net exports, which decrease because of the rise in the interest rate. This situation, in which expansionary fiscal policy does not lead to a rise in output, is frequently referred to as a case of **complete crowding out**.[2]

[2]When the demand for money is affected by the interest rate—the usual case in which the *LM* curve slopes upward but is not vertical—some crowding out occurs. The rightward shift of the *IS* curve also raises the interest rate, which causes investment spending and net exports to fall somewhat. However, as indicated in Figure S3-6, the rise in the interest rate is not sufficient to reduce investment spending and net exports to the point where aggregate output does not increase. Thus expansionary fiscal policy increases aggregate output, and only partial crowding out occurs.

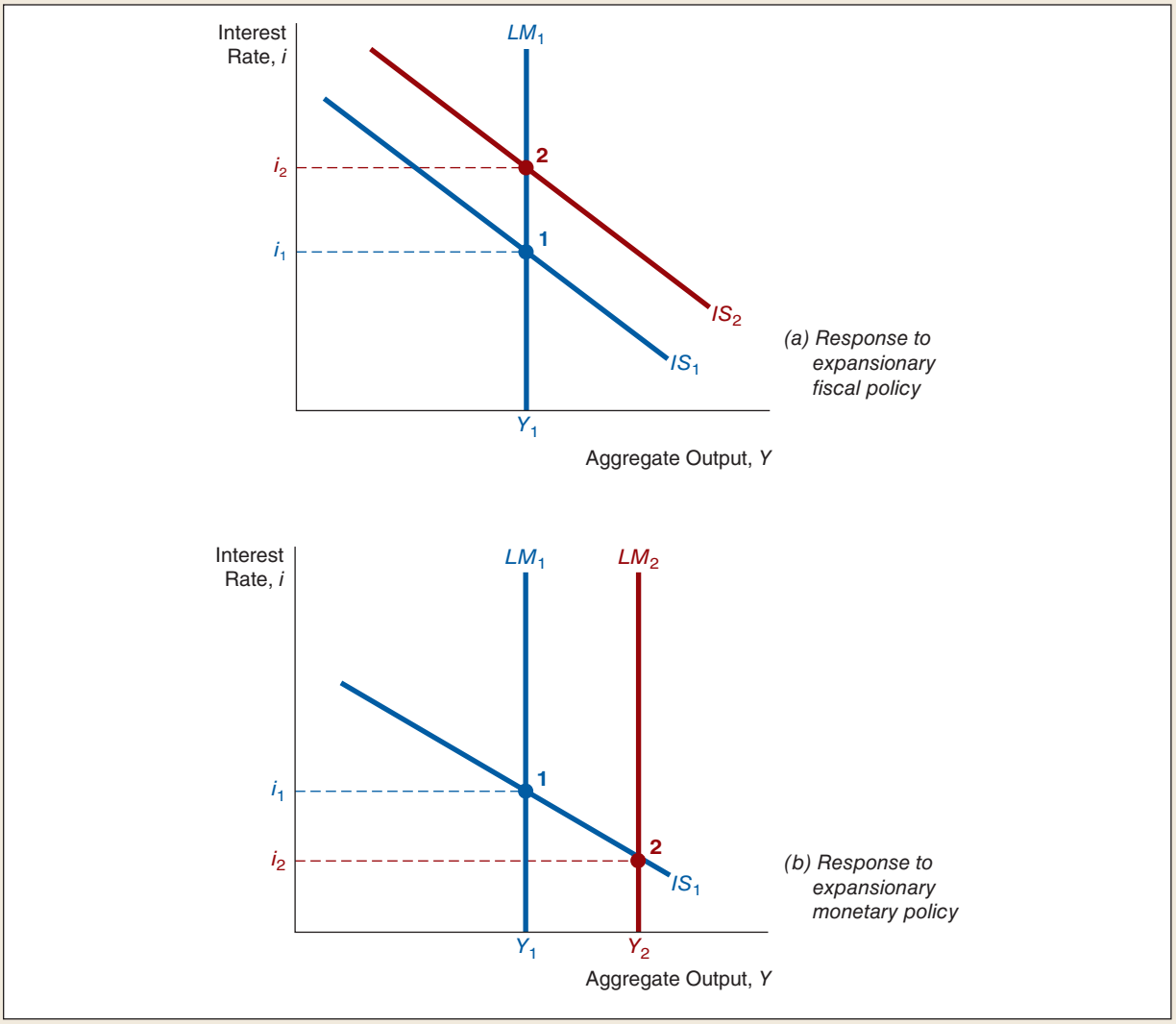


FIGURE S3-7 Effectiveness of Monetary and Fiscal Policy When Money Demand Is Unaffected by the Interest Rate

When the demand for money is unaffected by the interest rate, the *LM* curve is vertical. In panel (a), an expansionary fiscal policy (an increase in government spending or a cut in taxes) shifts the *IS* curve from IS_1 to IS_2 and leaves aggregate output unchanged at Y_1. In panel (b), an increase in the money supply shifts the *LM* curve from LM_1 to LM_2 and raises aggregate output from Y_1 to Y_2. Therefore, monetary policy is effective, but fiscal policy is not.

Panel (b) shows what happens when the Bank of Canada tries to eliminate high unemployment through an expansionary monetary policy (an increase in the real money supply, *M/P*). Here, the *LM* curve shifts to the right from LM_1 to LM_2 because, at each interest rate, output must rise so that the quantity of money demanded will rise to match the increase in the money supply. Aggregate output rises from Y_1 to Y_2

(the economy moves from point 1 to point 2), and expansionary monetary policy does affect aggregate output in this case.

We conclude from the analysis in Figure S3-7 that if the demand for money is unaffected by changes in the interest rate (money demand is interest-inelastic), monetary policy is effective but fiscal policy is not. An even more general conclusion can be reached: ***The less interest-sensitive money demand is, the more effective monetary policy is relative to fiscal policy.***[3]

Because the interest sensitivity of money demand is important to policymakers' decisions regarding the use of monetary or fiscal policy to influence economic activity, the subject has been studied extensively by economists and has been the focus of many debates. Findings on the interest sensitivity of money demand are discussed in Chapter 20.

APPLICATION Targeting Money Supply Versus Interest Rates

In the 1970s and early 1980s, central banks in many countries pursued a strategy of *monetary targeting*—that is, they used their policy tools to try to achieve a money supply equal to a target value. However, with the breakdown of the stable relationship between the money supply and economic activity, many of these central banks abandoned monetary targeting in the 1980s and pursued interest-rate targeting instead. The *ISLM* model gives us important information about which variable a central bank should target, and we can apply it to explain why central banks have abandoned monetary targeting and switched to interest-rate targeting.[4]

As we saw in Chapter 17, when the Bank of Canada attempts to hit a reserve aggregate or a money supply target, it cannot at the same time pursue an interest-rate target; it can hit one target or the other but not both. Consequently, it needs to know which of these two targets will produce more effective control of aggregate output.

In contrast to the textbook world that you have been inhabiting, in which the *IS* and *LM* curves are assumed to be fixed, the real world is one of great uncertainty. In the real world, *IS* and *LM* curves shift because of unanticipated changes in autonomous spending and money demand. To understand whether the Bank of Canada should use a money supply target or an interest-rate target, we need to look at two cases: one in which uncertainty about the *IS* curve is far greater than uncertainty about the *LM* curve, and another in which uncertainty about the *LM* curve is far greater than uncertainty about the *IS* curve.

The *ISLM* diagram in Figure S3-8 illustrates the outcomes of the two targeting strategies when the *IS* curve is unstable and uncertain, so that it fluctuates around its expected value of IS^* from IS' to IS'', and the *LM* curve is stable and certain, so that it stays at LM^*. Because the central bank knows that the expected position of the *IS* curve is at IS^* and desires aggregate output of Y^*, it will set its interest-rate target at i^*. This policy is represented by the line labelled "Interest-Rate Target" in Figure S3-8.

[3]This result and many others in this chapter can be obtained more directly by using algebra. An algebraic treatment of the *ISLM* model is given in an appendix to this chapter.

[4]The classic paper on this topic is William Poole, "The Optimal Choice of Monetary Policy Instruments in a Simple Macro Model," *Quarterly Journal of Economics* 84 (1970): 192–216. A less mathematical version of his analysis, far more accessible to students, is contained in William Poole, "Rules of Thumb for Guiding Monetary Policy," in *Open Market Policies and Operating Procedures: Staff Studies* (Washington, DC: Board of Governors of the Federal Reserve System, 1971).

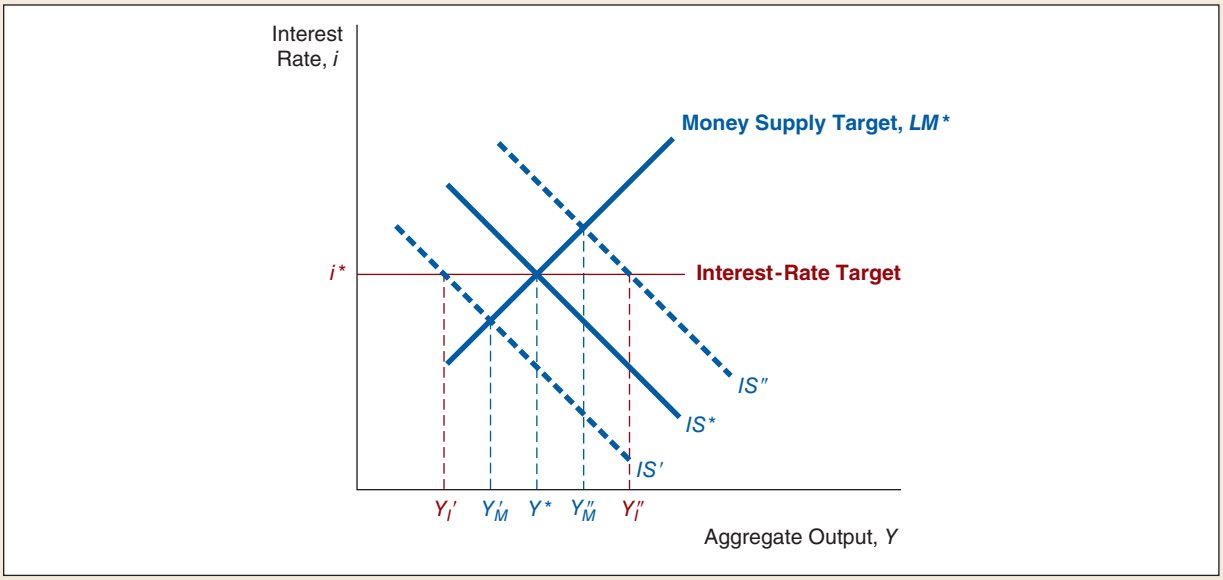


FIGURE S3-8 Money Supply and Interest-Rate Targets When the *IS* Curve Is Unstable and the *LM* Curve Is Stable

The unstable *IS* curve fluctuates between *IS′* and *IS″*. The money supply target produces smaller fluctuations in output (Y_M' to Y_M'') than the interest-rate target (Y_I' to Y_I''). Therefore, the money supply target is preferred over the interest-rate target.

How can the central bank keep the interest rate at its target level of *i**? Recall from Chapter 17 that when the interest rate differs from *i**, the Bank of Canada can hit the interest-rate target by buying and selling bonds. When the *IS* curve shifts out to *IS″*, the interest rate rises above *i**, with the money supply unchanged. To counter this rise in interest rates, the central bank will need to buy bonds just until their price is driven back up enough so that the interest rate comes back down to *i**. (As we saw in Chapter 15, the result of these open market purchases is a rise in both the monetary base and the money supply until the *LM* curve shifts to the right to intersect the *IS″* curve at *i**—not shown in the diagram, for simplicity.) When the interest rate is below *i**, the central bank needs to sell bonds in order to lower their price and raise the interest rate back up to *i**. (These open market sales reduce the monetary base and the money supply until the *LM* curve shifts to the left to intersect the *IS* curve at *i**—again, not shown in the diagram.) Pursuit of the interest-rate target causes aggregate output to fluctuate between Y_I' and Y_I'' in Figure S3-8.

If, instead, the Bank of Canada decides to pursue a money supply target, it will set the money supply target such that the resulting *LM* curve *LM** intersects the *IS** curve at the desired output level of *Y**. This policy is represented by the line labelled "Money Supply Target." Because the money supply is held constant and so the *LM* curve is fixed at *LM**, aggregate output fluctuates between Y_M' and Y_M'' under the money supply target policy.

As you can see in Figure S3-8, the money supply target leads to smaller output fluctuations around the desired level than the interest-rate target. Given a money supply target, a rightward shift of the *IS* curve (to *IS″*, for example) will cause the interest rate to rise, and this rise in the interest rate will lead to a lower level of investment spending and

net exports and hence to a smaller increase in aggregate output than would occur under an interest-rate target. Because smaller output fluctuations are desirable, the conclusion is that ***if the IS curve is more unstable than the LM curve, a money supply target is preferred over an interest-rate target***.

Figure S3-9 illustrates the outcomes of the two targeting strategies when the *IS* curve is stable and the *LM* curve is unstable (because of unanticipated changes in money demand). Again, the interest-rate and money supply targets are set such that the expected level of aggregate output is equal to the desired level Y^*. Because the *LM* curve is now unstable, it fluctuates between LM' and LM'' even when the money supply is fixed, causing aggregate output to fluctuate between Y'_M and Y''_M.

The interest-rate target, by contrast, is not affected by uncertainty about the *LM* curve, because it is determined by the Bank of Canada's adjustment of the money supply whenever the interest rate tries to depart from i^*. When the interest rate begins to rise above i^* because of an increase in money demand, the central bank just buys bonds, driving up their price and bringing the interest rate back down to i^*. The result of these open market purchases is a rise in both the monetary base and the money supply. Similarly, if the interest rate falls below i^*, the central bank sells bonds to lower their price and to raise the interest rate back to i^*, thereby causing declines in the monetary base and the money supply. The only effect of the fluctuating *LM* curve, then, is that the money supply fluctuates more as a result of the interest-rate target policy. With the interest-rate target, output is exactly at the desired level, with no fluctuations.

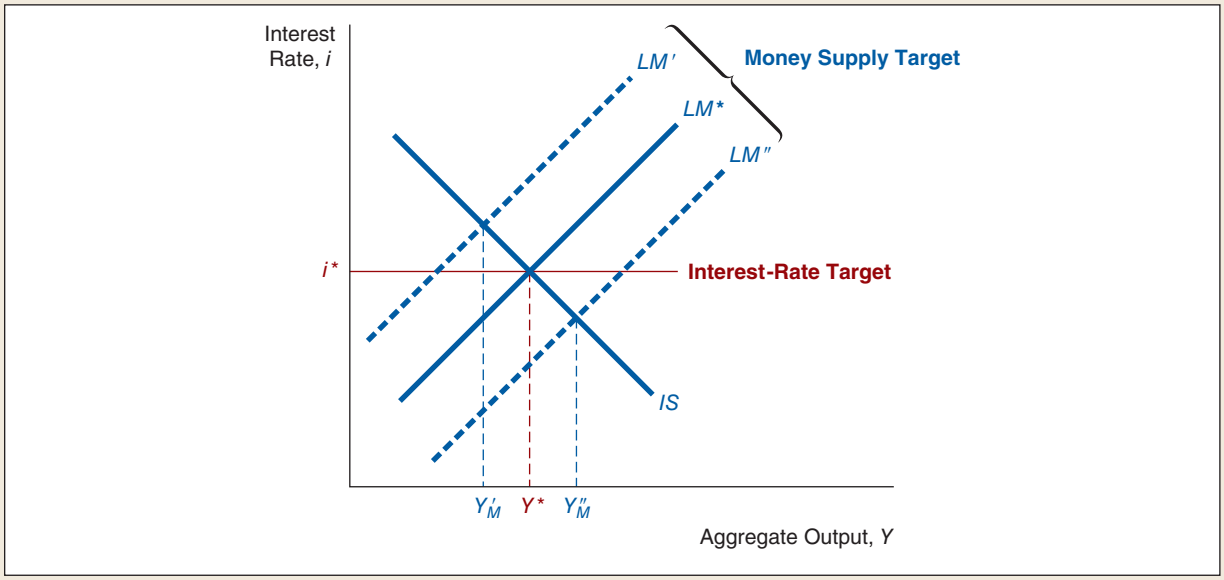


FIGURE S3-9 Money Supply and Interest-Rate Targets When the *LM* Curve Is Unstable and the *IS* Curve Is Stable

The unstable *LM* curve fluctuates between LM' and LM''. The money supply target then produces bigger fluctuations in output (Y'_M to Y''_M) than the interest-rate target (which leaves output fixed at Y^*). Therefore, in this case, the interest-rate target is preferred.

Because smaller output fluctuations are desirable, the conclusion from Figure S3-9 is that ***if the LM curve is more unstable than the IS curve, an interest-rate target is preferred over a monetary target***.

We can now see why many central banks decided to abandon monetary targeting for interest-rate targeting in the 1980s. With the rapid proliferation of new financial instruments whose presence affected the demand for money, money demand (which is embodied in the *LM* curve) became highly unstable in many countries. Thus central banks in these countries recognized that their economies were more likely to be in the situation depicted in Figure S3-9 and decided that they would be better off with an interest-rate target than a money supply target.[5] ◆

S3.6 *ISLM* MODEL IN THE LONG RUN

So far in our *ISLM* analysis, we have assumed that the price level is fixed so that nominal values and real values are the same. This is a reasonable assumption for the short run, but in the long run the price level does change. To see what happens in the *ISLM* model in the long run, we make use of the concept of *natural rate of output*, more often referred to as *potential output*: the rate of output at which the price level has no tendency to rise or fall. Potential output (denoted by Y^P) is the level to which the economy settles in the long run. When output is above potential output, the booming economy will cause prices to rise; when output is below potential output, the slack in the economy will cause prices to fall.

Because we now want to examine what happens when the price level changes, we can no longer assume that real and nominal values are the same. The spending variables that affect the *IS* curve (consumer expenditure, investment spending, government spending, and net exports) describe the demand for goods and services *in real terms*; they describe the physical quantities of goods that people want to buy. Because these quantities do not change when the price level changes, an alteration in the price level has no effect on the *IS* curve, which describes *in real terms* the combinations of the interest rate and aggregate output that satisfy goods market equilibrium.

Figure S3-10 shows what happens in the *ISLM* model when output rises above potential output, which is marked by a vertical line at Y^P. Suppose that the *IS* and *LM* curves initially intersect at point 1, where output $Y = Y^P$. Panel (a) examines what happens to output and interest rates when the money supply increases. As we saw in Figure S3-3, the increase in the money supply causes the *LM* curve to shift rightward

[5]It is important to recognize, however, that the crucial factor in deciding which target is preferable is the *relative* instability of the *IS* and *LM* curves. Although the *LM* curve has been unstable recently, the evidence supporting a stable *IS* curve is also weak. Instability in the money demand function does not automatically mean that money supply targets should be abandoned for interest-rate targets. Furthermore, the analysis so far has been conducted under the assumption that the price level is fixed. More realistically, when the price level is allowed to change, there is uncertainty about expected inflation, and so the case for an interest-rate target is less strong. As we learned in Chapters 4 and 5, the interest rate that is more relevant to investment decisions is not the nominal interest rate but the real interest rate (the nominal interest rate minus expected inflation). Hence, when expected inflation rises, the real interest rate falls at each given nominal interest rate, and investment and net exports rise, shifting the *IS* curve to the right. Similarly, a fall in expected inflation raises the real interest rate at each given nominal interest rate, lowers investment and net exports, and shifts the *IS* curve to the left. In the real world, expected inflation undergoes large fluctuations, so the *IS* curve in Figure S3-9 will undergo substantial fluctuations, making it less likely that the interest-rate target is preferable to the money supply target.

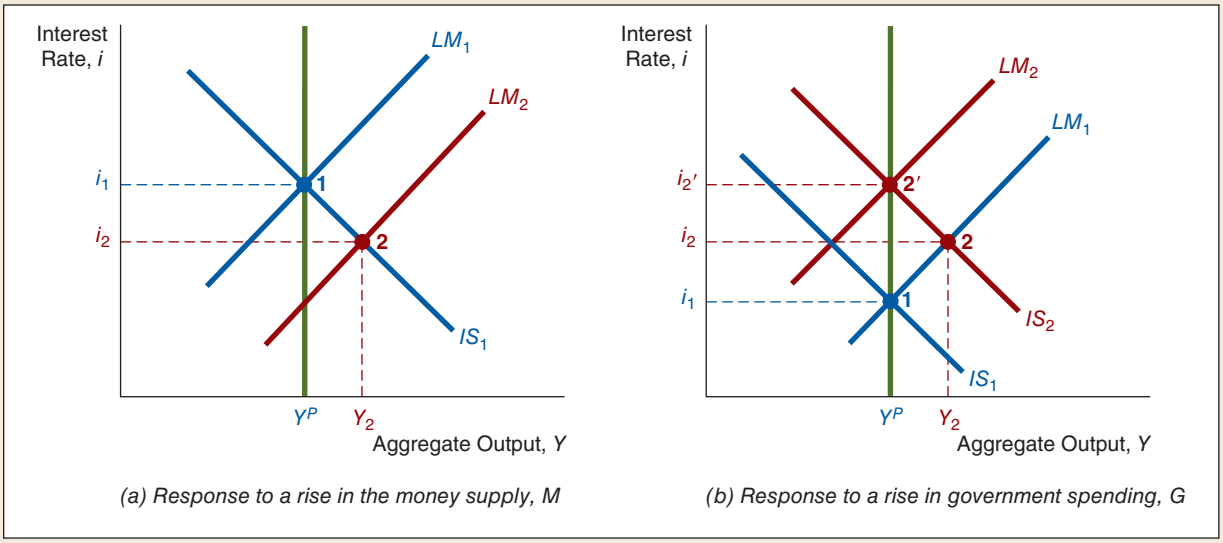


FIGURE S3-10 ***ISLM* Model in the Long Run**

In panel (a), an increase in the money supply causes the *LM* curve to shift rightward to LM_2, and the equilibrium moves to point 2, where the interest rate falls to i_2 and output rises to Y_2. Because output at Y_2 is above potential output Y^P, the price level rises, the real money supply falls, and the *LM* curve shifts back to LM_1; the economy has returned to the original equilibrium at point 1. In panel (b), an increase in government spending shifts the *IS* curve to the right to IS_2, and the economy moves to point 2, where the interest rate has risen to i_2 and output has risen to Y_2. Because output at Y_2 is above potential output Y^P, the price level begins to rise, real money balances *M/P* begin to fall, and the *LM* curve shifts to the left to LM_2. The long-run equilibrium at point 2′ has an even higher interest rate at $i_{2'}$, and output has returned to Y^P.

to LM_2, and the equilibrium moves to point 2 (the intersection of IS_1 and LM_2), where the interest rate falls to i_2 and output rises to Y_2. However, as we can see in panel (a), the level of output at Y_2 is greater than potential output at Y^P, and so the price level begins to rise.

In contrast to the *IS* curve, which is unaffected by a rise in the price level, the *LM* curve is affected by the price level rise because the liquidity preference theory states that the demand for money *in real terms* depends on real income and interest rates. This theory makes sense because money is valued in terms of what it can buy. However, the money supply reported in dollars by the media is not the money supply in real terms; it is a nominal quantity. As the price level rises, the quantity of money *in real terms* falls, and the effect on the *LM* curve is identical to a fall in the nominal money supply with the price level fixed. The lower value of the real money supply creates an excess demand for money, causing the interest rate to rise at any given level of aggregate output, and the *LM* curve shifts back to the left. As long as the level of output exceeds potential output, the price level will continue to rise, shifting the *LM* curve to the left, until finally output is back at potential output Y^P. This occurs when the *LM* curve has returned to LM_1, where real money balances *M/P* have returned to the original level and the economy has returned to the original equilibrium at point 1. In the long run, the level of output and interest rates are unchanged by the expansion in the money supply.

The fact that an increase in the money supply leaves output and interest rates unchanged in the long run is referred to as **long-run monetary neutrality**. The only result of the increase in the money supply is a higher price level, which has increased proportionally to the increase in the money supply, so that real money balances *M/P* are unchanged.

Panel (b) looks at what happens to output and interest rates under expansionary fiscal policy, such as an increase in government spending. As we saw earlier, the increase in government spending shifts the *IS* curve to the right to IS_2, and in the short run the economy moves to point 2 (the intersection of IS_2 and LM_1), where the interest rate has risen to i_2 and output has risen to Y_2. Because output at Y_2 is above potential output Y^P, the price level begins to rise, real money balances *M/P* begin to fall, and the *LM* curve shifts to the left. Only when the *LM* curve has shifted to LM_2 and the equilibrium has reached point 2′, where output is again at potential output Y^P, does the price level stop rising and the *LM* curve come to rest. The resulting long-run equilibrium at point 2′ is characterized by an even higher interest rate at $i_{2'}$, and output has not risen from Y^P. Indeed, what has occurred in the long run is complete crowding out; the rise in the price level, which has shifted the *LM* curve to LM_2, has caused the interest rate to rise to $i_{2'}$, causing investment and net exports to fall enough to offset the increased government spending completely. We have discovered that even though complete crowding out does not occur in the short run in the *ISLM* model (unless the *LM* curve is vertical), it does occur in the long run.

We can conclude from our study of the effects of monetary and fiscal policy on the *ISLM* model that ***although monetary and fiscal policy can affect output in the short run, neither affects output in the long run.*** Clearly, in order to use monetary and fiscal policy effectively to raise output, policymakers must know where output is relative to potential output and how soon the long run occurs.

SUMMARY

1. The *ISLM* model uses the *IS* and *LM* curves to determine aggregate output and the interest rate at a fixed price level. The *IS* curve traces out the combinations of the interest rate and aggregate output for which the goods market is in equilibrium, and the *LM* curve traces out the combinations for which the market for money is in equilibrium. The *IS* curve slopes downward because higher interest rates lower planned investment spending and net exports and so lower equilibrium output. The *LM* curve slopes upward because higher aggregate output raises the demand for money and so raises the equilibrium interest rate.
2. The intersection point of the *IS* and *LM* curves gives the simultaneous output and interest rates at which both the goods market and the market for money are in equilibrium. At any other level of interest rates and output, at least one of the markets will be out of equilibrium, and economic forces will move the economy toward the general equilibrium point at the intersection of the *IS* and *LM* curves.
3. The *LM* curve is shifted to the right by a rise in the money supply or an autonomous fall in money demand; it is shifted to the left by a fall in the money supply or an autonomous rise in money demand.
4. A rise in the money supply raises equilibrium output but lowers the equilibrium interest rate. Expansionary fiscal policy (a rise in government spending or a fall in taxes) raises equilibrium output but, in contrast to expansionary monetary policy, also raises the interest rate.
5. The less interest-sensitive money demand is, the more effective monetary policy is relative to fiscal policy.

6. The *ISLM* model leads us to the following conclusion about the conduct of monetary policy: When the *IS* curve is more unstable than the *LM* curve, a money supply target provides smaller output fluctuations than an interest-rate target does and so is the preferred target; when the *LM* curve is more unstable than the *IS* curve, an interest-rate target leads to smaller output fluctuations than a money supply target and so is the preferred target.
7. By examining the effects of expansionary monetary or fiscal policy on the *ISLM* model, we can conclude that although monetary and fiscal policy can affect output in the short run, neither affects output in the long run.

MyLab Economics Make the grade with MyLab Economics: Select end-of-chapter questions can be found on MyLab Economics. You can practise them as often as you want, and most feature step-by-step guided instructions to help you find the right answer.

QUESTIONS

1. Using a supply and demand diagram for the market for money, show what happens to the equilibrium level of the interest rate as aggregate output falls. What does this result imply about the slope of the *LM* curve?
2. "If the point describing a combination of the interest rate and aggregate output is not on either the *IS* curve or the *LM* curve, the economy will have no tendency to head toward the intersection of the two curves." Is this statement true, false, or uncertain? Explain your answer.
3. What happens to the position of the *LM* curve if the Bank of Canada decides that it will decrease the money supply to fight inflation and if, at the same time, the demand for money falls?
4. "An excess demand for money resulting from a rise in the demand for money can be eliminated only by a rise in the interest rate." Is this statement true, false, or uncertain? Explain your answer.

In Problems 5–11, demonstrate your answers with an ISLM *diagram.*

5. In late 1969, the Federal Reserve in the United States reduced the money supply while the government raised taxes. What do you think should have happened to interest rates and aggregate output?
6. Suppose the Bank of Canada wants to keep interest rates from rising when the government sharply increases military spending. How can the Bank do this?
7. Evidence indicates that the demand for money has become quite unstable in recent years. Why is this finding important to Bank of Canada policymakers?
8. "As the price level rises, the equilibrium level of output determined in the *ISLM* model also rises." Is this statement true, false, or uncertain? Explain your answer.
9. Predict what will happen to interest rates and aggregate output if a stock market crash causes autonomous consumer expenditure to fall.
10. Predict what will happen to interest rates and aggregate output if an autonomous export boom occurs.
11. If a series of defaults in the bond market makes bonds riskier and, as a result, the demand for money rises, predict what will happen to interest rates and aggregate output.

APPENDIX TO SUPPLEMENTAL CHAPTER 3

Algebra of the *ISLM* Model

The use of algebra to analyze the *ISLM* model enables us to obtain many of the results in Supplemental Chapter 3 very quickly.

BASIC CLOSED-ECONOMY *ISLM* MODEL

For an economy with no exports or imports, the goods market can be described by the following equations:

Consumption function:	$C = \overline{C} + mpc\,(Y - T)$	(1)
Investment function:	$I = \overline{I} - di$	(2)
Taxes:	$T = \overline{T}$	(3)
Government spending:	$G = \overline{G}$	(4)
Goods market equilibrium condition:	$Y = Y^{ad} = C + I + G$	(5)

The money market can be described by the following equations:

Money demand function:	$M^d = \overline{M}^d + eY - mi$	(6)
Money supply:	$M^s = \overline{M}$	(7)
Money market equilibrium condition:	$M^d = M^s$	(8)

The uppercase terms are the variables of the model. $\overline{G}$, $\overline{T}$, and $\overline{M}$ represent the policy variables that are set exogenously (outside the model); $\overline{C}$, $\overline{I}$, and $\overline{M}^d$ are the autonomous components of consumer expenditure, investment spending, and money demand, respectively, which are also determined exogenously. Except for the interest rate i, the lowercase terms are the parameters, or the givens of the model, and all are assumed to be positive. Output (Y) and interest rate (i) are the endogenous variables whose values are determined in the *ISLM* model. The definitions of these variables and parameters are as follows:

C = consumer spending
I = investment spending
$G = \overline{G}$ = government spending
Y = output
$T = \overline{T}$ = taxes
M^d = money demand
$M^s = \overline{M}$ = money supply
i = interest rate
$\overline{C}$ = autonomous consumer spending

d = interest sensitivity of investment spending
$\bar{I}$ = autonomous investment spending related to business confidence
$\overline{M}^d$ = autonomous money demand
e = income sensitivity of money demand
m = interest sensitivity of money demand
mpc = marginal propensity to consume

IS and *LM* Curves

Substituting for C, I, and G in the goods market equilibrium condition and then solving for Y, we obtain the *IS* curve

$$Y = \frac{1}{1 - mpc}(\overline{C} + \bar{I} - mpc\,\overline{T} + \overline{G} - di) \tag{9}$$

Solving for i using Equations 6, 7, and 8, we obtain the *LM* curve

$$i = \frac{\overline{M}^d - \overline{M} + eY}{m} \tag{10}$$

Solution of the Model

The solution of the model gives us the point of intersection of the *IS* and *LM* curves. To find this intersection point, we solve for Y and i simultaneously, using Equations 9 and 10, as follows:

$$Y = \frac{1}{1 - mpc + de/m}\left(\overline{C} + \bar{I} - mpc\,\overline{T} + \overline{G} - \frac{d\overline{M}^d}{m} + \frac{d\overline{M}}{m}\right) \tag{11}$$

$$i = \frac{1}{m(1 - mpc) + d}[e(\overline{C} + \bar{I} - mpc\,\overline{T} + \overline{G}) + \overline{M}^d(1 - mpc) - \overline{M}(1 - mpc)] \tag{12}$$

Implications

The conclusions reached through these algebraic solutions are the same as those reached in the supplemental chapter; for example:

1. Because all of the coefficients are positive, Equation 11 indicates that a rise in $\overline{C}$, $\bar{I}$, $\overline{G}$, and $\overline{M}$ leads to a rise in Y and that a rise in $\overline{T}$ or $\overline{M}^d$ leads to a fall in Y.
2. Equation 12 indicates that a rise in $\overline{C}$, $\bar{I}$, $\overline{G}$, and $\overline{M}^d$ leads to a rise in i and that a rise in $\overline{M}$ or $\overline{T}$ leads to a fall in i.
3. As m, the interest sensitivity of money demand, increases, the multiplier term

$$\frac{1}{1 - mpc + de/m}$$

increases, and so fiscal policy ($\overline{G}$, $\overline{T}$) has more effect on output; conversely, the term multiplying $\overline{M}$,

$$\frac{d}{m}\left(\frac{1}{1 - mpc + de/m}\right) = \frac{d}{m(1 - mpc) + de}$$

decreases, so monetary policy has less effect on output.

4. By similar reasoning, as d, the interest sensitivity of investment spending, increases, monetary policy has more effect on output and fiscal policy has less effect on output.

OPEN-ECONOMY *ISLM* MODEL

To make the basic *ISLM* model into an open-economy model, we need to include net exports in the goods market equilibrium condition so that Equation 5 becomes Equation 5′

$$Y = Y^{ad} = C + I + G + NX \tag{5}'$$

As suggested by the discussion in Chapter 21, the net exports and exchange rate relations can be written as follows:

$$NX = \overline{NX} - hE \tag{13}$$
$$E = \overline{E} + ji \tag{14}$$

where
NX = net exports
$\overline{NX}$ = autonomous net exports
h = exchange rate sensitivity of net exports
E = exchange rate (value of domestic currency)
$\overline{E}$ = autonomous exchange rate
j = interest sensitivity of exchange rate

Substituting for net exports in the goods market equilibrium condition (Equation 5′) using the net exports and exchange rate relations and then solving for Y as in the basic model, we obtain the open-economy *IS* curve

$$Y = \frac{1}{1 - mpc}[\overline{C} + \overline{I} - mpc\,\overline{T} + \overline{G} + \overline{NX} - h\overline{E} - (d + hj)i] \tag{15}$$

The *LM* curve is the same as in the basic model, and so the solutions for Y and i are as follows:

$$Y = \frac{1}{1 - mpc + (d + hj)e/m} \times \left(\overline{C} + \overline{I} - mpc\,\overline{T} + \overline{G} - \frac{d + hj}{m}\overline{M}^d + \frac{d + hj}{m}\overline{M} + \overline{NX} - h\overline{E}\right) \tag{16}$$

$$i = \frac{1}{m(1 - mpc) + (d + hj)e} \times [e(\overline{C} + \overline{I} - mpc\,\overline{T} + \overline{G} + \overline{NX} - h\overline{E}) + \overline{M}^d(1 - mpc) - \overline{M}(1 - mpc)] \tag{17}$$

IMPLICATIONS

1. As indicated by the *IS* curve given in Equation 15, the inclusion of net exports in aggregate demand provides an additional reason for the negative relationship between Y and i (the downward slope of the *IS* curve). This additional reason for the negative relationship between Y and i is represented by hj in the term $-(d + hj)i$.
2. Equations 16 and 17 indicate that all of the results generated by the basic model hold true for the open-economy model as well.
3. Equation 16 indicates that a rise in $\overline{NX}$ leads to a rise in Y, and an autonomous rise in the value of the domestic currency $\overline{E}$ leads to a decline in Y.
4. Equation 17 indicates that a rise in $\overline{NX}$ leads to a rise in i, and a rise in $\overline{E}$ leads to a decline in i.

PART 2

Financial Markets

Crisis and Response: Credit Market Turmoil and the Stock Market Crashes of October 2008 and March 2020

The financial crisis that started during the summer of 2007 in the United States began to snowball as the value of mortgage-backed securities on financial institutions' balance sheets plummeted. A "flight to quality" drove three-month Treasury bill rates down to almost zero, rates not seen since the Great Depression of the 1930s. Credit spreads—an indicator of risk—shot through the roof, with the gap between Eurodollar and Treasury bill rates (the TED spread) going from around 40 basis points (0.40 percentage points) before the financial crisis began to over 450 basis points in mid-October, the highest value in its history. After earlier sharp declines, the stock market crashed further, with the week beginning on October 6, 2008, showing the worst weekly decline in U.S. history.

When Covid-19 became a worldwide pandemic, the three-month Treasury bill rate in the United States again fell to near zero; the stock market crashed, falling by more than 35% from February 19, 2020, to March 23, 2020; and the TED spread rose from around 10 basis points in February to over 140 basis points by the end of March.

The recent financial crisis and the impact of the coronavirus pandemic illustrate how volatile financial markets can be. This volatility can hit financial consumers directly, putting jobs in jeopardy and leading to difficulty in getting loans, falling home values, and declining retirement account values. How can policy respond to disruptions in financial markets? We begin addressing this question by examining the inner workings of financial markets, particularly interest rate dynamics. Chapter 4 explains what an interest rate is, as well as the relationships among interest rates, bond prices, and returns. Chapter 5 examines how the overall level of interest rates is determined. In Chapter 6, we extend the analysis of the bond market to explain changes in credit spreads and the relationship of long-term to short-term interest rates. Chapter 7 looks at the role of expectations in the stock market and explores what drives stock prices.

4 The Meaning of Interest Rates

Learning Objectives

4.1 Calculate the present value of future cash flows and the yield to maturity on the four types of credit market instruments.

4.2 Recognize the distinctions among yield to maturity, current yield, rate of return, and rate of capital gain.

4.3 Interpret the distinction between real and nominal interest rates.

Preview

Interest rates are among the most closely watched variables in the economy. Their movements are reported almost daily by the news media because they directly affect our everyday lives and have important consequences for the health of the economy. They influence personal decisions such as whether to consume or save, whether to buy a house, and whether to purchase bonds or put funds into a savings account. Interest rates also affect the economic decisions of businesses and households, such as whether to use their funds to invest in new equipment for factories or to save rather than spend their money.

Before we can go on with the study of money, banking, and financial markets, we must understand exactly what the phrase *interest rates* means. In this chapter, we see that a concept known as the *yield to maturity* is the most accurate measure of interest rates; the yield to maturity is what economists mean when they use the term *interest rate*. We'll discuss how the yield to maturity is measured. We'll also see that a bond's interest rate does not necessarily indicate how good an investment the bond is, because what the bond earns (its rate of return) does not necessarily equal its interest rate. Finally, we'll explore the distinction between real interest rates, which are adjusted for inflation, and nominal interest rates, which are not.

Although learning definitions is not always the most exciting of pursuits, it is important to read carefully and understand the concepts presented in this chapter. Not only are they used continually throughout the remainder of this text, but a firm grasp of these terms will give you a clearer understanding of the role that interest rates play in your life as well as in the general economy.

4.1 MEASURING INTEREST RATES

LO 4.1 Calculate the present value of future cash flows and the yield to maturity on the four types of credit market instruments.

Different debt instruments have very different streams of cash payments (known as **cash flows**) to the holder, with very different timing. Thus we first need to understand how we can compare the value of one kind of debt instrument with the value of another before we see how interest rates are measured. To do this, we make use of the concept of *present value*.

Present Value

The concept of **present value** (or **present discounted value**) is based on the commonsense notion that a dollar paid to you one year from now is less valuable than a dollar paid to you today. This notion is true because you can deposit a dollar today in a savings account that earns interest and have more than a dollar in one year. Economists use a more formal definition, as explained in this section.

Let's look at the simplest kind of debt instrument, which we will call a **simple loan**. In this loan, the lender provides the borrower with an amount of funds (called the *principal*) that must be repaid to the lender at the *maturity date*, along with an additional payment for the interest. For example, if you made your friend, Jane, a simple loan of $100 for one year, you would require her to repay the principal of $100 in one year's time, along with an additional payment for interest—say, $10. In the case of a simple loan like this one, the interest payment divided by the amount of the loan is a natural and sensible way to measure the interest rate. This measure of the so-called *simple interest rate*, i, is

$$i = \frac{\$10}{\$100} = 0.10 = 10\%$$

If you make this $100 loan, at the end of the year you will have $110, which can be rewritten as

$$\$100 \times (1 + 0.10) = \$110$$

If you then lent out the $110 at the same interest rate, at the end of the second year you would have

$$\$110 \times (1 + 0.10) = \$121$$

or, equivalently,

$$\$100 \times (1 + 0.10) \times (1 + 0.10) = \$100 \times (1 + 0.10)^2 = \$121$$

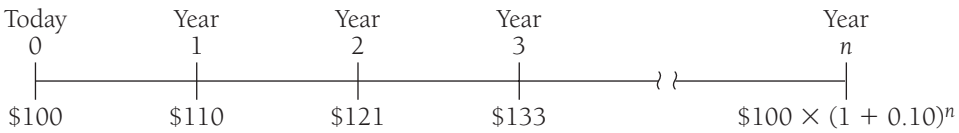
Continuing with the loan again, at the end of the third year you would have

$$\$121 \times (1 + 0.10) = \$100 \times (1 + 0.10)^3 = \$133$$

Generalizing, we can see that at the end of n years, your $100 would turn into

$$\$100 \times (1 + i)^n$$

The amounts you would have at the end of each year by making the $100 loan today can be seen in the following timeline:



This timeline clearly indicates that you are just as happy having $100 today as you will be having $110 a year from now (of course, as long as you are sure that Jane will pay you back). You are also just as happy having $100 today as you will be having $121 two years from now, or $133 three years from now, or $100 \times (1 + 0.10)^n$ dollars n years from now. The timeline indicates that we can also work backward from future

amounts to the present. For example, $\$133 = \$100 \times (1 + 0.10)^3$ three years from now is worth $100 today, so

$$\$100 = \frac{\$133}{(1 + 0.10)^3}$$

The process of calculating today's value of dollars received in the future, as we have done above, is called *discounting the future*. This is why the concept of present value is also referred to as *present discounted value*. We can generalize this process by writing today's (present) value of $100 as *PV* and the future cash flow (payment) of $133 as *CF* and then replacing 0.10 (the 10% interest rate) by *i*. This leads to the following formula:

$$PV = \frac{CF}{(1 + i)^n} \tag{1}$$

Intuitively, Equation 1 tells us that if you are promised $1 of cash flow, for certain, 10 years from now, this dollar would not be as valuable to you as $1 is today because, if you had the $1 today, you could invest it and end up with more than $1 in 10 years.

The concept of present value is extremely useful, because it allows us to figure out today's value (price) of a credit (debt) market instrument at a given simple interest rate *i* by just adding up the individual present values of all the future payments received. This information enables us to compare the values of two or more instruments that have very different timing of their payments.

APPLICATION Simple Present Value

What is the present value of $250 to be paid in two years if the interest rate is 15%?

Solution

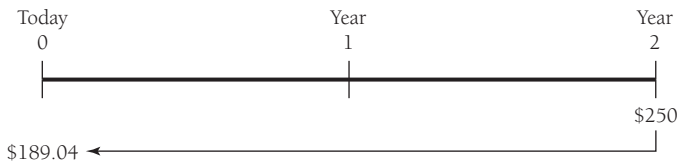
The present value would be $189.04. We find this value by using Equation 1,

$$PV = \frac{CF}{(1 + i)^n}$$

where
CF = cash flow in two years = 250
i = annual interest rate = 0.15
n = number of years = 2

Thus,

$$PV = \frac{\$250}{(1 + 0.15)^2} = \frac{\$250}{1.3225} = \$189.04$$



APPLICATION How Much Is That Jackpot Worth?

Assume that you just hit the $20 million jackpot in a lottery, which promises you a payment of $1 million every year for the next 20 years. You are clearly excited, but have you really won $20 million?

Solution

No, not in the present value sense. In today's dollars, that $20 million is worth a lot less. If we assume an interest rate of 10% as in the earlier examples, the first payment of $1 million is clearly worth $1 million today, but the second payment next year is worth only $1 million$/(1 + 0.10) = \$909\,090$, a lot less than $1 million. The following year, the payment is worth $1 million$/(1 + 0.10)^2 = \$826\,446$ in today's dollars, and so on. When you add up all these amounts, they total $9.4 million. You are still pretty excited (who wouldn't be?), but because you understand the concept of present value, you recognize that you are the victim of false advertising. In present-value terms, you didn't really win $20 million but instead won less than half that amount. (And that's before taxes!) ◆

Four Types of Credit Market Instruments

In terms of the timing of their cash flow payments, there are four basic types of credit market instruments.

1. A simple loan, which we have already discussed, in which the lender provides the borrower with an amount of funds that must be repaid to the lender at the maturity date, along with an additional payment for the interest. Many money market instruments are of this type—for example, short-term commercial loans to businesses.
2. A **fixed-payment loan** (also called a **fully amortized loan**) in which the lender provides the borrower with an amount of funds that the borrower must repay by making the same payment, consisting of part of the principal and interest, every period (such as a month) for a set number of years. For example, if you borrow $1000, a fixed-payment loan might require you to pay $126 every year for 25 years. Installment loans (such as auto loans) and mortgages are frequently of the fixed-payment type.
3. A **coupon bond** pays the owner of the bond a fixed interest payment (coupon payment) every year until the maturity date, when a specified final amount (**face value** or **par value**) is repaid. (The coupon payment is so named because in the past, the bondholder obtained payment by clipping a coupon off the bond and sending it to the bond issuer, who then sent the payment to the holder. Today, it is no longer necessary to send in coupons to receive payments.) A coupon bond with $1000 face value, for example, might pay you a coupon payment of $100 per year for 10 years and then repay you the face value amount of $1000 at the maturity date. (The face value of a bond is usually in $1000 increments.)

   A coupon bond is identified by four pieces of information: First is the bond's face value; second is the corporation or government agency that issues the bond; third is the maturity date of the bond; and fourth is the bond's **coupon rate**, which is the dollar amount of the yearly coupon payment expressed as a percentage of the face value of the bond. In our example, the coupon bond has a yearly coupon payment of $100 and a face value of $1000. The coupon rate is then $\$100/\$1000 = 0.10$, or 10%. Capital market instruments such as Government of Canada bonds and notes and corporate bonds are examples of coupon bonds.

4. A **discount bond** (also called a **zero-coupon bond**) is bought at a price below its face value (at a discount), and the face value is repaid at the maturity date. Unlike a coupon bond, a discount bond does not make any interest payments; it just pays off the face value. For example, a one-year discount bond with a face value of $1000 might be bought for $900; in a year's time, the owner would be repaid the face value of $1000. Canadian government Treasury bills and long-term zero-coupon bonds are examples of discount bonds.

These four types of instruments make payments at different times: Simple loans and discount bonds make payments only at their maturity dates, whereas fixed-payment loans and coupon bonds make payments periodically until maturity. How can you decide which of these instruments will provide you with the most income? They all seem so different because they make payments at different times. To solve this problem, we use the concept of present value, explained earlier, to provide us with a procedure for measuring interest rates on these different types of instruments.

Yield to Maturity

Of the several common ways of calculating interest rates, the most important is the **yield to maturity**, which is the interest rate that equates the present value of cash flow payments received from a debt instrument with its value today.[1] Because the concept behind the calculation of the yield to maturity makes good economic sense, economists consider it the most accurate measure of interest rates.

To better understand the yield to maturity, we now look at how it is calculated for the four types of credit market instruments. In all of these examples, the key to understanding the calculation of the yield to maturity is realizing that we are equating today's value of the debt instrument with the present value of all of its future cash flow payments.

Simple Loan

Using the concept of present value, the yield to maturity on a simple loan is easy to calculate. For the one-year loan we discussed, today's value is $100, and the payments in one year's time would be $110 (the repayment of $100 plus the interest payment of $10). We can use this information to solve for the yield to maturity i by recognizing that the present value of the future payments must equal today's value of the loan.

APPLICATION Yield to Maturity on a Simple Loan

If Pete borrows $100 from his sister and next year she wants $110 back from him, what is the yield to maturity on this loan?

Solution

The yield to maturity on the loan is 10%.

$$PV = \frac{CF}{(1+i)^n}$$

where
PV = amount borrowed = $100
CF = cash flow in one year = $110
n = number of years = 1

[1] In other contexts, the yield to maturity is also called the *internal rate of return*.

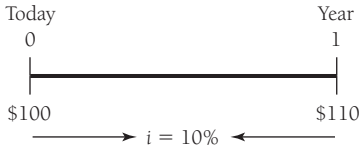
Thus,

$$\$100 = \frac{\$110}{(1+i)}$$

$$(1+i)\$100 = \$110$$

$$(1+i) = \frac{\$110}{\$100}$$

$$i = 1.10 - 1 = 0.10 = 10\%$$



◆

This calculation of the yield to maturity should look familiar because it equals the interest payment of $10 divided by the loan amount of $100; that is, it equals the simple interest rate on the loan. An important point to recognize is that ***for simple loans, the simple interest rate equals the yield to maturity***. Hence the same term i is used to denote both the yield to maturity and the simple interest rate.

Fixed-Payment Loan Recall that this type of loan has the same cash flow payment every period throughout the life of the loan. On a fixed-rate mortgage, for example, the borrower makes the same payment to the bank every month until the maturity date, at which time the loan will be completely paid off. To calculate the yield to maturity for a fixed-payment loan, we follow the same strategy that we used for the simple loan—we equate today's value of the loan with its present value. Because the fixed-payment loan involves more than one cash flow payment, the present value of the fixed-payment loan is calculated (using Equation 1) as the sum of the present values of all cash flow payments.

In the case of our earlier example, the loan is $1000 and the yearly payment is $126 for the next 25 years. The present value (PV) is calculated as follows: At the end of one year, there is a $126 payment with a PV of $\$126/(1+i)$; at the end of two years, there is another $126 payment with a PV of $\$126/(1+i)^2$; and so on until, at the end of the twenty-fifth year, the last payment of $126 with a PV of $\$126/(1+i)^{25}$ is made. Setting today's value of the loan ($1000) equal to the sum of the present values of all the yearly payments gives us

$$\$1000 = \frac{\$126}{1+i} + \frac{\$126}{(1+i)^2} + \frac{\$126}{(1+i)^3} + \ \ldots\ + \frac{\$126}{(1+i)^{25}}$$

More generally, for any fixed-payment loan,

$$LV = \frac{FP}{(1+i)} + \frac{FP}{(1+i)^2} + \frac{FP}{(1+i)^3} + \ \ldots\ + \frac{FP}{(1+i)^n} \qquad (2)$$

where
LV = loan value
FP = fixed yearly payment
n = number of years until maturity

For a fixed-payment loan, the loan value, the fixed yearly payment, and the number of years until maturity are known quantities, and only the yield to maturity is not. So we can solve this equation for the yield to maturity i. Because this calculation is not easy, many financial calculators have programs that enable you to find i given the loan's numbers for LV, FP, and n. For example, in the case of a 25-year \$1000 loan with yearly payments of \$85.81, the yield to maturity that results from solving Equation 2 is 7%. Real estate brokers always have handy a financial calculator that can solve such equations so that they can immediately tell the prospective house buyer exactly what the yearly (or monthly) payments will be if the house purchase is financed by a mortgage.

APPLICATION Yield to Maturity and the Yearly Payment on a Fixed-Payment Loan

You decide to purchase a new home and need a \$100 000 mortgage. You take out a loan from the bank that has an interest rate of 7%. What is the yearly payment to the bank if you wish to pay off the loan in 20 years?

Solution

The yearly payment to the bank is \$9 439.29.

$$LV = \frac{FP}{(1+i)} + \frac{FP}{(1+i)^2} + \frac{FP}{(1+i)^3} + \ldots + \frac{FP}{(1+i)^n}$$

where
LV = loan value amount = 100 000
i = annual interest rate = 0.07
n = number of years = 20

Thus,

$$\$100\,000 = \frac{FP}{(1+0.07)} + \frac{FP}{(1+0.07)^2} + \frac{FP}{(1+0.07)^3} + \ldots + \frac{FP}{(1+0.07)^{20}}$$

To find the monthly payment for the loan using a financial calculator:

n = number of years = 20
PV = amount of the loan (LV) = 100 000
FV = amount of the loan after 20 years = 0
i = annual interest rate = 0.07

Then push the *PMT* button to get the fixed yearly payment (FP) = \$9 439.29. ◆

Coupon Bond To calculate the yield to maturity for a coupon bond, follow the same strategy used for the fixed-payment loan: Equate today's value of the bond with its present value. Because coupon bonds also have more than one cash flow payment, the present value of the bond is calculated as the sum of the present values of all the coupon payments plus the present value of the final payment of the face value of the bond.

The present value of a $1000-face-value bond with 10 years to maturity and yearly coupon payments of $100 (a 10% coupon rate) can be calculated as follows: At the end of one year, there is a $100 coupon payment with a *PV* of $\$100/(1+i)$; at the end of the second year, there is another $100 coupon payment with a *PV* of $\$100/(1+i)^2$; and so on until, at maturity, there is a $100 coupon payment with a *PV* of $\$100/(1+i)^{10}$ plus the repayment of the $1000 face value with a *PV* of $\$1000/(1+i)^{10}$. Setting today's value of the bond (its current price, denoted by *P*) equal to the sum of the present values of all the cash flow payments for the bond gives

$$P = \frac{\$100}{1+i} + \frac{\$100}{(1+i)^2} + \frac{\$100}{(1+i)^3} + \ \ldots\ + \frac{\$100}{(1+i)^{10}} + \frac{\$1000}{(1+i)^{10}}$$

More generally, for any coupon bond,[2]

$$P = \frac{C}{1+i} + \frac{C}{(1+i)^2} + \frac{C}{(1+i)^3} + \ \ldots\ + \frac{C}{(1+i)^n} + \frac{F}{(1+i)^n} \qquad (3)$$

where

P = price of the coupon bond
C = yearly coupon payment
F = face value of the bond
n = years to maturity date

In Equation 3, the coupon payment, the face value, the years to maturity, and the price of the bond are known quantities, and only the yield to maturity is not. Hence we can solve this equation for the yield to maturity *i*. As in the case of the fixed-payment loan, this calculation is not easy, so business-oriented software and financial calculators have built-in programs that solve the equation for you.

APPLICATION Yield to Maturity and Bond Price for a Coupon Bond

Find the price of a 10% coupon bond with a face value of $1000, a 12.25% yield to maturity, and eight years to maturity.

Solution

The price of the bond is $889.20. To solve using a financial calculator

n = years to maturity = 8
FV = face value of the bond (F) = 1000
i = yield to maturity = 12.25%
PMT = yearly coupon payments (C) = 100

Then push the *PV* button to get the price of the bond = $889.20.

[2]Most coupon bonds actually make coupon payments on a semiannual basis rather than once a year, as we assumed here. The effect on the calculations is very slight and will be ignored here.

Alternatively, you could solve for the yield to maturity given the bond price by entering \$889.20 for *PV* and pushing the *i* button to get a yield to maturity of 12.25%. ◆

Table 4-1 shows the yields to maturity calculated for several bond prices. Three interesting facts emerge:

1. When the coupon bond is priced at its face value, the yield to maturity equals the coupon rate.
2. The price of a coupon bond and the yield to maturity are negatively related; that is, as the yield to maturity rises, the price of the bond falls. As the yield to maturity falls, the price of the bond rises.
3. The yield to maturity is greater than the coupon rate when the bond price is below its face value and is less than the coupon rate when the bond price is above its face value.

These three facts are true for any coupon bond and are really not surprising if you think about the reasoning behind the calculation of the yield to maturity. When you put \$1000 in a bank account with an interest rate of 10%, you can take out \$100 every year, and you will be left with the \$1000 at the end of 10 years. This process is similar to buying the \$1000 bond with a 10% coupon rate analyzed in Table 4-1, which pays a \$100 coupon payment every year and then repays \$1000 at the end of 10 years. If the bond is purchased at the par value of \$1000, its yield to maturity must equal 10%, which is also equal to the coupon rate of 10%. The same reasoning applied to any coupon bond demonstrates that if the coupon bond is purchased at its par value, the yield to maturity and the coupon rate must be equal.

It is a straightforward process to show that the bond price and the yield to maturity are negatively correlated. As *i*, the yield to maturity, increases, all denominators in the bond price formula (Equation 3) must necessarily increase, because the rise in *i* lowers the present value of all future cash flow payments for this bond. Hence a rise in the interest rate, as measured by the yield to maturity, means that the price of the bond must fall. Another way to explain why the bond price falls when the interest rate rises is to consider that a higher interest rate implies that the future coupon payments and final payment are worth less when discounted back to the present; hence the price of the bond must be lower.

The third fact, that the yield to maturity is greater than the coupon rate when the bond price is below its par value, follows directly from facts 1 and 2. When the yield to maturity equals the coupon rate, the bond price is at the face value; when the yield

TABLE 4-1 Yields to Maturity on a 10%-Coupon-Rate Bond Maturing in 10 Years (Face Value = \$1000)

Price of Bond (\$)	Yield to Maturity (%)
1200	7.13
1100	8.48
1000	10.00
900	11.75
800	13.81

to maturity rises above the coupon rate, the bond price necessarily falls and so must be below the face value of the bond.

One special case of a coupon bond is worth discussing here because its yield to maturity is particularly easy to calculate. This bond is called a **consol** or a **perpetuity**; it is a perpetual bond with no maturity date and no repayment of principal that makes fixed coupon payments of $C forever. Consols were first sold by the British Treasury during the Napoleonic wars and are still traded today; they are quite rare, however, in North American capital markets. The formula in Equation 3 for the price of the consol P_c simplifies to the following:[3]

$$P_c = \frac{C}{i_c} \tag{4}$$

where P_c = price of the perpetuity (consol)
C = yearly payment
i_c = yield to maturity of the perpetuity (consol)

One nice feature of perpetuities is that you can immediately see that as i_c increases, the price of the bond falls. For example, if a perpetuity pays $100 per year forever and the interest rate is 10%, its price will be $1000 = $100/0.10. If the interest rate rises to 20%, its price will fall to $500 = $100/0.20. We can rewrite this formula as

$$i_c = \frac{C}{P_c} \tag{5}$$

APPLICATION Yield to Maturity on a Perpetuity

What is the yield to maturity on a bond that has a price of $2000 and pays $100 of interest annually, forever?

Solution

The yield to maturity is 5%.

$$i_c = \frac{C}{P_c}$$

[3]The bond price formula for a consol is

$$P = \frac{C}{1+i} + \frac{C}{(1+i)^2} + \frac{C}{(1+i)^3} + \ldots$$

which can be written as

$$P = C(x + x^2 + x^3 + \ldots)$$

in which $x = 1/(1+i)$. The formula for an infinite sum is

$$1 + x + x^2 + x^3 + \ldots = \frac{1}{1-x} \text{ for } x < 1$$

and so,

$$P = C\left(\frac{1}{1-x} - 1\right) = C\left[\frac{1}{1-1/(1+i)} - 1\right]$$

which by suitable algebraic manipulation becomes

$$P = C\left(\frac{1+i}{i} - \frac{i}{i}\right) = \frac{C}{i}$$

where C = yearly payment = \$100
P_c = price of perpetuity (consol) = \$2000

Thus,

$$i_c = \frac{\$100}{\$2000}$$

$$i_c = 0.05 = 5\%$$ ◆

The formula given in Equation 5, which describes the calculation of the yield to maturity for a perpetuity, also provides a useful approximation for the yield to maturity on coupon bonds. When a coupon bond has a long term to maturity (say, 20 years or more), it is very much like a perpetuity, which pays coupon payments forever. This is because the cash flows more than 20 years in the future have such small present discounted values that the value of a long-term coupon bond is very close to the value of a perpetuity with the same coupon rate. Thus i_c in Equation 5 will be very close to the yield to maturity for any long-term bond. For this reason, i_c, the yearly coupon payment divided by the price of the security, has been given the name **current yield** and is frequently used as an approximation to describe interest rates on long-term bonds.

Discount Bond The yield-to-maturity calculation for a discount bond is similar to that for a simple loan. Let's consider a discount bond such as a one-year Canadian Treasury bill that pays a face value of \$1000 in one year's time but today has a price of \$900.

APPLICATION Yield to Maturity on a Discount Bond

What is the yield to maturity on a one-year, \$1000 Treasury bill with a current price of \$900?

Solution

The yield to maturity is 11.1%.
Using the present value formula,

$$PV = \frac{CF}{(1+i)^n}$$

and recognizing that the present value (PV) is the current price of \$900, the cash flow in one year is \$1000, and the number of years is 1, we can write:

$$\$900 = \frac{\$1000}{1+i}$$

Solving for i, we get

$$(1+i) \times \$900 = \$1000$$

$$\$900 + \$900i = \$1000$$

$$\$900i = \$1000 - \$900$$

$$i = \frac{\$1000 - \$900}{\$900} = 0.111 = 11.1\%$$ ◆

As we just saw in the preceding application, the yield to maturity for a one-year discount bond equals the increase in price over the year, \$1000 − \$900, divided by the initial price, \$900. Hence, more generally, for any one-year discount bond, the yield to maturity can be written as

$$i = \frac{F - P}{P} \tag{6}$$

where F = face value of the discount bond
P = current price of the discount bond

In other words, the yield to maturity equals the increase in price over the year $F - P$ divided by the initial price P. In normal circumstances, investors earn positive returns from holding these securities and so they sell at a discount, meaning that the current price of the bond is below the face value. Therefore, $F - P$ should be positive, and the yield to maturity should be positive as well. However, this is not always the case, as extraordinary events in Japan and elsewhere have indicated (see the Global box).

An important feature of this equation is that it indicates that, for a discount bond, the yield to maturity is negatively related to the current bond price. This is the same conclusion that we reached for a coupon bond. Equation 6 shows that a rise in the bond price—say, from \$900 to \$950—means that the bond will have a smaller increase in its price at maturity and so the yield to maturity will fall, from 11.1% to 5.3% in our example. Similarly, a fall in the yield to maturity means that the current price of the discount bond has risen.

Summary The concept of present value tells you that a dollar in the future is not as valuable to you as a dollar today because you can earn interest on a dollar you have today. Specifically, a dollar received n years from now is worth only $\$1/(1+i)^n$ today. The present value of a set of future cash flow payments on a debt instrument equals the sum of the present values of each of the future payments. The yield to maturity for an instrument is the interest rate that equates the present value of the future payments on that instrument to its value today. Because the procedure for calculating the yield to maturity is based on sound economic principles, the yield to maturity is the measure that economists think most accurately describes the interest rate.

Our calculations of the yield to maturity for a variety of bonds reveal the important fact that ***current bond prices and interest rates are negatively related: When the interest rate rises, the price of the bond falls, and vice versa***.

4.2 THE DISTINCTION BETWEEN INTEREST RATES AND RETURNS

LO 4.2 Recognize the distinctions among yield to maturity, current yield, rate of return, and rate of capital gain.

Many people think that the interest rate on a bond tells them all they need to know about how well off they are as a result of owning it. If Irving the Investor thinks he is well off when he owns a long-term bond yielding a 10% interest rate, and the interest rate then rises to 20%, he will have a rude awakening: As we will shortly see, if he has to sell the bond, Irving will lose his shirt! How well a person does financially by holding a bond or any other security over a particular time period is accurately measured by the security's **return**, or, in more precise terminology, the **rate of return**. We will

Global Negative Interest Rates? Japan First, Then the United States, Then Europe

We normally assume that the yield to maturity must always be positive. A negative yield to maturity would imply that you are willing to pay more for a bond today than you will receive for it in the future (as our formula for yield to maturity on a discount bond demonstrates). A negative yield to maturity therefore seems like an impossibility because you would do better by holding cash that has the same value in the future as it does today.

Events in Japan in the late 1990s and then in the United States during the 2008 global financial crisis and finally in Europe in recent years have demonstrated that this reasoning is not quite correct. In November 1998, the yield to maturity on Japanese six-month Treasury bills became negative, at −0.004%. In September 2008, the yield to maturity on three-month U.S. T-bills fell very slightly below zero for a very brief period. Interest rates that banks are paid on deposits they keep at their central banks became negative first in Sweden in July 2009, followed by Denmark in July 2012, the Eurozone in June 2014, Switzerland in December 2014, and Japan in January 2016. Negative interest rates have rarely occurred in the past. How could this happen in recent years?

As we will see in Chapter 5, a dearth of investment opportunities and very low inflation can drive interest rates to low levels, but these two factors can't explain the negative yield to maturity. The answer is that despite negative interest rates, large investors and banks found it more convenient to hold Treasury bills or keep their funds as deposits at the central bank because they are stored electronically. For that reason, investors and banks were willing to accept negative interest rates even though in pure monetary terms the investors would be better off holding cash.

use the concept of *return* continually throughout this book: Understanding this concept will make the material presented later in the book easier to follow.

For any security, the rate of return is defined as the amount of each payment to the owner plus the change in the security's value, expressed as a fraction of its purchase price. To make this definition clearer, let us see what the return would look like for a $1000-face-value coupon bond with a coupon rate of 10% that is bought for $1000, held for one year, and then sold for $1200. The payments to the owner are the yearly coupon payments of $100, and the change in the bond's value is $1200 − $1000 = $200. Adding these values together and expressing them as a fraction of the purchase price of $1000 gives us the one-year holding-period return for this bond:

$$\frac{\$100 + \$200}{\$1000} = \frac{\$300}{\$1000} = 0.30 = 30\%$$

You may have noticed something quite surprising about the return that we just calculated: It equals 30%, yet as Table 4-1 indicates, initially the yield to maturity was only 10%. This discrepancy demonstrates that ***the return on a bond will not necessarily equal the yield to maturity on that bond***. We now see that the distinction between interest rate and return can be important, although for many securities the two may be closely related.

More generally, the return on a bond held from time t to time $t + 1$ can be written as

$$R = \frac{C + P_{t+1} - P_t}{P_t} \tag{7}$$

where R = return from holding the bond from time t to time $t + 1$
C = coupon payment
P_t = price of the bond at time t
P_{t+1} = price of the bond at time $t + 1$

A convenient way to rewrite the return formula in Equation 7 is to recognize that it can be split into two separate terms:

$$R = \frac{C}{P_t} + \frac{P_{t+1} - P_t}{P_t}$$

The first term is the current yield i_c (the coupon payment over the purchase price):

$$\frac{C}{P_t} = i_c$$

The second term is the **rate of capital gain**, or the change in the bond's price relative to the initial price:

$$\frac{P_{t+1} - P_t}{P_t} = g$$

where g is the rate of capital gain. Equation 7 can then be rewritten as

$$R = i_c + g \tag{8}$$

which shows that the return on a bond is the current yield i_c plus the rate of capital gain g. This rewritten formula illustrates the point we just discovered: Even for a bond for which the current yield i_c is an accurate measure of the yield to maturity, the return can differ substantially from the interest rate. Returns will differ substantially from the interest rate if the price of the bond experiences sizable fluctuations that produce substantial capital gains or losses.

To explore this point even further, let's look at what happens to the returns on bonds of different maturities when interest rates rise. Table 4-2 calculates the one-year returns, using Equation 8 above, on several 10%-coupon-rate bonds, all purchased at par, when interest rates on all these bonds rise from 10% to 20%. Several key findings from this table are generally true of all bonds:

- The only bonds whose returns will equal their initial yields to maturity are those whose times to maturity are the same as their holding periods (see, for example, the last bond in Table 4-2).
- A rise in interest rates is associated with a fall in bond prices, resulting in capital losses on bonds whose terms to maturity are longer than their holding periods.
- The more distant a bond's maturity date, the greater the size of the percentage price change associated with an interest-rate change.
- The more distant a bond's maturity date, the lower the rate of return that occurs as a result of an increase in the interest rate.
- Even though a bond may have a substantial initial interest rate, its return can turn out to be negative if interest rates rise.

TABLE 4-2 One-Year Returns on Different-Maturity 10%-Coupon-Rate Bonds When Interest Rates Rise from 10% to 20%

(1) Years to Maturity When Bond Is Purchased	(2) Initial Current Yield (%)	(3) Initial Price ($)	(4) Price Next Year* ($)	(5) Rate of Capital Gain (%)	(6) Rate of Return [col (2) + col (5)] (%)
30	10	1000	503	−49.7	−39.7
20	10	1000	516	−48.4	−38.4
10	10	1000	597	−40.3	−30.3
5	10	1000	741	−25.9	−15.9
2	10	1000	917	−8.3	+1.7
1	10	1000	1000	0.0	+10.0

*Calculated with a financial calculator, using Equation 3.

At first, students are frequently puzzled (as was poor Irving the Investor) that a rise in interest rates can mean that a bond has been a poor investment. The trick to understanding this fact is to recognize Irving has already purchased the bond, so a rise in the interest rate means that the price of the bond Irving is holding falls and he experiences a capital loss. If this loss is large enough, the bond can be a poor investment indeed. For example, we see in Table 4-2 that the bond that has 30 years to maturity when purchased has a capital loss of 49.7% when the interest rate rises from 10% to 20%. This loss is so large that it exceeds the current yield of 10%, resulting in a negative return (loss) of –39.7%. If Irving does not sell the bond, his capital loss is often referred to as a "paper loss." This is a loss nonetheless because if Irving had not bought this bond and had instead put his money in the bank, he would now be able to buy more bonds at their lower price than he presently owns.

Maturity and the Volatility of Bond Returns: Interest-Rate Risk

The finding that the prices of longer-maturity bonds respond more dramatically to changes in interest rates helps explain an important fact about the behaviour of bond markets: ***Prices and returns for long-term bonds are more volatile than those for shorter-term bonds.*** Price changes of +20% and −20% within a year, with corresponding variations in returns, are common for bonds that are more than 20 years away from maturity.

We now see that changes in interest rates make investments in long-term bonds quite risky. Indeed, the risk level associated with an asset's return that results from interest-rate changes is so important that it has been given a special name, **interest-rate risk**.[4] Dealing with interest-rate risk is a major concern of managers of financial institutions and of investors, as we will see in later chapters.

[4]Interest-rate risk can be quantitatively measured using the concept of *duration*. This concept and its calculation are discussed in Online Appendix 4.1.

Although long-term debt instruments have substantial interest-rate risk, short-term debt instruments do not. Indeed, bonds with a maturity term that is as short as the holding period have no interest-rate risk.[5] We see this for the coupon bond at the bottom of Table 4-2, which has no uncertainty about the rate of return because it equals the yield to maturity, which is known at the time the bond is purchased. The key to understanding why there is no interest-rate risk for *any* bond whose time to maturity matches the holding period is to recognize that (in this case) the price at the end of the holding period is already fixed at the face value. A change in interest rates can then have no effect on the price at the end of the holding period for these bonds, and the return will therefore be equal to the yield to maturity, which is known at the time the bond is purchased.[6]

Summary

The return on a bond, which tells you how good an investment it has been over the holding period, is equal to the yield to maturity in only one special case—when the holding period and the term to maturity of the bond are identical. Bonds whose terms to maturity are longer than their holding periods are subject to interest-rate risk: Changes in interest rates lead to capital gains and losses that produce substantial differences between the return and the yield to maturity known at the time the bond is purchased. Interest-rate risk is especially important

[5]The statement that there is no interest-rate risk for any bond whose time to maturity matches the holding period is literally true only for discount (zero-coupon) bonds that make no intermediate cash payments before the holding period is over. A coupon bond that makes an intermediate cash payment before the holding period is over requires this payment to be reinvested. Because the interest rate at which this payment can be reinvested is uncertain, some uncertainty exists about the return on this coupon bond even when the time to maturity equals the holding period. However, the riskiness of the return on a coupon bond from reinvesting the coupon payments is typically quite small, so a coupon bond with a time to maturity equal to its holding period still has very little risk.

[6]In this text we assume that holding periods on all short-term bonds are equal to the term to maturity, and thus the bonds are not subject to interest-rate risk. However, if the term to maturity of the bond is shorter than an investor's holding period, the investor is exposed to a type of interest-rate risk called *reinvestment risk*. Reinvestment risk occurs because the proceeds from the short-term bond need to be reinvested at a future interest rate that is uncertain.

To understand reinvestment risk, suppose that Irving the Investor has a holding period of two years and decides to purchase a $1000, one-year bond at face value and then another one at the end of the first year. If the initial interest rate is 10%, Irving will have $1100 at the end of the first year. If the interest rate rises to 20%, as in Table 4-2, Irving will find that buying $1100 worth of another one-year bond will leave him at the end of the second year with $\$1100 \times (1 + 0.20) = \$1320$. Thus Irving's two-year return will be $(\$1320 - \$1000)/1000 = 0.32 = 32\%$, which equals 14.9% at an annual rate. In this case, Irving has earned more by buying the one-year bonds than he would have if he had initially purchased a two-year bond with an interest rate of 10%. Thus, when Irving has a holding period that is longer than the term to maturity of the bonds he purchases, he benefits from a rise in interest rates. Conversely, if interest rates fall to 5%, Irving will have only $1155 at the end of two years: $\$1100 \times (1 + 0.05)$. His two-year return will be $(\$1555 - \$1000)/1000 = 0.155 = 15.5\%$, which is 7.2% at an annual rate. With a holding period greater than the term to maturity of the bond, Irving now loses from a decline in interest rates.

We have seen that when the holding period is longer than the term to maturity of a bond, the return is uncertain because the future interest rate when reinvestment occurs is also uncertain—in short, there is reinvestment risk. We also see that if the holding period is longer than the term to maturity of the bond, the investor benefits from a rise in interest rates and is hurt by a fall in interest rates.

for long-term bonds, on which capital gains and losses can be substantial. This is why long-term bonds are not considered safe assets with a sure return over short holding periods.

4.3 THE DISTINCTION BETWEEN REAL AND NOMINAL INTEREST RATES

LO 4.3 Interpret the distinction between real and nominal interest rates.

So far in our discussion of interest rates, we have ignored the effects of inflation on the cost of borrowing. What we up to this point have been calling the interest rate makes no allowance for inflation, and it is more precisely referred to as the **nominal interest rate**. We must distinguish the nominal interest rate from the **real interest rate**, which is the interest rate that is adjusted by subtracting expected changes in the price level (inflation) so that it more accurately reflects the true cost of borrowing. This interest rate is more precisely referred to as the *ex ante real interest rate* because it is adjusted for *expected* changes in the price level. The ex ante real interest rate is very important to economic decisions, and typically it is what economists mean when they make reference to the "real" interest rate. The interest rate that is adjusted for actual changes in the price level is called the *ex post real interest rate*. It describes how well a lender has done in real terms *after the fact*.

The real interest rate is more accurately defined from the *Fisher equation*, named for Irving Fisher, one of the great monetary economists of the twentieth century. The Fisher equation states that the nominal interest rate i equals the real interest rate r plus the expected rate of inflation π^e:[7]

$$i = r + \pi^e \tag{9}$$

Rearranging terms, we find that the real interest rate equals the nominal interest rate minus the expected inflation rate:

$$r = i - \pi^e \tag{10}$$

To see why this definition makes sense, let's first consider a situation in which you have made a simple one-year loan with a 5% interest rate $(i = 5\%)$, and you expect the price level to rise by 3% over the course of the year $(\pi^e = 3\%)$. As a result of making the loan, at the end of the year you expect to have 2% more in **real terms**—that

[7]A more precise formulation of the Fisher equation is

$$i = r + \pi^e + (r \times \pi^e)$$

because

$$1 + i = (1 + r)(1 + \pi^e) = 1 + r + \pi^e + (r \times \pi^e)$$

and subtracting 1 from both sides gives us the first equation. For small values of r and π^e, the term $r \times \pi^e$ is so small that we ignore it, as in the text.

is, in terms of real goods and services you can buy. In this case, the interest rate you expect to earn in terms of real goods and services is 2%:

$$r = 5\% - 3\% = 2\%$$

as indicated by Equation 10.[8]

APPLICATION Calculating Real Interest Rates

What is the real interest rate if the nominal interest rate is 8% and the expected inflation rate is 10% over the course of a year?

Solution

The real interest rate is −2%. Although you will be receiving 8% more dollars at the end of the year, you will be paying 10% more for goods. The result is that you will be able to buy 2% fewer goods at the end of the year, and you will be 2% worse off in real terms. Mathematically,

$$r = i - \pi^e$$

where i = nominal interest rate = 0.08
π^e = expected inflation rate = 0.10

Thus,

$$r = 0.08 - 0.10 = -0.02 = -2\%$$ ◆

As a lender, you are clearly less eager to make a loan in this case, because in terms of real goods and services you have actually earned a negative interest rate of 2%. By contrast, as a borrower, you fare quite well because at the end of the year, the amounts you will have to pay back will be worth 2% less in terms of goods and services—you as the borrower will be ahead by 2% in real terms. ***When the real interest rate is low, there are greater incentives to borrow and fewer incentives to lend.***

A similar distinction can be made between nominal returns and real returns. Nominal returns, which do not allow for inflation, are what we have been referring to as simply "returns." When inflation is subtracted from a nominal return, we have the real

[8]Because most interest income in Canada is subject to federal income taxes, the true earnings in real terms from holding a debt instrument are not reflected by the real interest rate defined by the Fisher equation but rather by the *after-tax real interest rate*, which equals the nominal interest rate *after income tax payments have been subtracted*, minus the expected inflation rate. For a person facing a 30% tax rate, the after-tax interest rate earned on a bond yielding 10% is only 7% because 30% of the interest income must be paid to Revenue Canada. Thus the after-tax real interest rate on this bond when expected inflation is 5% equals 2% (= 7% − 5%).

More generally, the after-tax real interest rate can be expressed as

$$i(1 - \tau) - \pi^e$$

where τ = the income tax rate.

As the formula indicates, after-tax real interest rates are always below the real interest rate defined by the Fisher equation. For a further discussion of measures of after-tax real interest rates, see Frederic S. Mishkin, "The Real Interest Rate: An Empirical Investigation," *Carnegie-Rochester Conference Series on Public Policy* 15 (1981): 151–200.

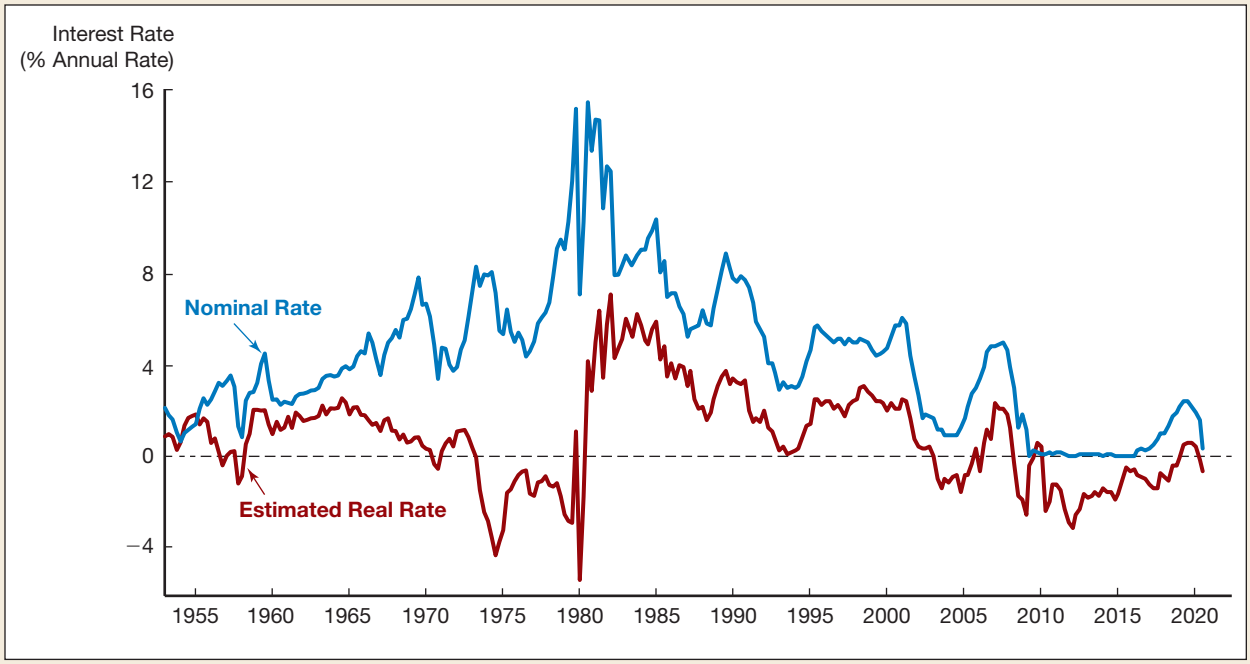


FIGURE 4-1 **Real and Nominal Interest Rates (Three-Month Treasury Bill), 1953–2020**

Nominal and real interest rates often do not move together. When U.S. nominal rates were high in the 1970s, real rates were actually extremely low—often negative.

Sources: Nominal rates from Federal Reserve Bank of St. Louis FRED database: fred.stlouisfed.org/series/TB3MS and fred.stlouisfed.org/series/CPIAUCSL. The real rate is constructed using the procedure outlined in Frederic S. Mishkin, "The Real Interest Rate: An Empirical Investigation," *Carnegie-Rochester Conference Series on Public Policy* 15 (1981): 151–200. This procedure involves estimating expected inflation as a function of past interest rates, inflation, and time trends and then subtracting the expected inflation measure from the nominal interest rate.

return, which indicates the amount of extra goods and services that we can purchase as a result of holding the security.

The distinction between real and nominal interest rates is important because the real interest rate, which reflects the real cost of borrowing, is likely to be a better indicator of the incentives to borrow and lend. It appears to be a better guide to how people will respond to what is happening in credit markets. Figure 4-1, which presents estimates from 1953 to 2020 of the real and nominal interest rates on three-month U.S. Treasury bills, shows us that nominal and real rates usually move together but do not always do so. (This is also true for nominal and real interest rates in Canada and the rest of the world.) In particular, when nominal rates were high in the 1970s, real rates were actually extremely low—often negative. By the standard of nominal interest rates, you would have thought that credit market conditions were tight during this period because it was expensive to borrow. However, the estimates of the real rates indicate that you would have been mistaken. In real terms, the cost of borrowing was actually quite low.

SUMMARY

1. The yield to maturity, which is the measure that most accurately reflects the interest rate, is the interest rate that equates the present value of future payments of a debt instrument with the instrument's value today. Application of this principle reveals that bond prices and interest rates are negatively correlated: When the interest rate rises, the price of the bond must fall, and vice versa.
2. The return on a security, which tells you how well you have done by holding the security over a stated period of time, can differ substantially from the interest rate as measured by the yield to maturity. Long-term bond prices experience substantial fluctuations when interest rates change and thus bear interest-rate risk. The resulting capital gains and losses can be large, which is why long-term bonds are not considered safe assets with a sure return.
3. The real interest rate is defined as the nominal interest rate minus the expected rate of inflation. It is both a better measure of the incentives to borrow and lend and a more accurate indicator of the tightness of credit market conditions than is the nominal interest rate.

MyLab Economics Make the grade with MyLab Economics: Select end-of-chapter questions can be found on MyLab Economics. You can practise them as often as you want, and most feature step-by-step guided instructions to help you find the right answer.

QUESTIONS

1. Would a dollar tomorrow be worth more to you today when the interest rate is 20% or when it is 10%?
2. Explain which information you would need to take into consideration when deciding to receive $5000 today or $5500 one year from today.
3. To help pay for university, you have just taken out a $1000 government loan that makes you pay $126 per year for 25 years. However, you don't have to start making these payments until you graduate from college two years from now. Why is the yield to maturity necessarily less than 12%? (This is the yield to maturity on a normal $1000 fixed-payment loan on which you pay $126 per year for 25 years.)
4. Do bondholders fare better when the yield to maturity increases or when it decreases? Why?
5. Suppose today you buy a coupon bond that you plan to sell one year later. Which part of the rate of return formula incorporates future changes into the bond's price? *Note:* Check Equations 7 and 8 in this chapter.
6. If mortgage rates rise from 5% to 10%, but the expected rate of increase in housing prices rises from 2% to 9%, are people more or less likely to buy houses?
7. When is the current yield a good approximation of the yield to maturity?
8. Why would a government choose to issue a perpetuity, which requires payments forever, instead of a terminal loan, such as a fixed-payment loan, discount bond, or coupon bond?
9. Under what conditions will a discount bond have a negative nominal interest rate? Is it possible for a coupon bond or a perpetuity to have a negative nominal interest rate?
10. True or False: With a discount bond, the return on the bond is equal to the rate of capital gain.
11. If interest rates decline, which would you rather be holding, long-term bonds or short-term bonds? Why? Which type of bond has the greater interest-rate risk?

12. Interest rates were lower in the mid-1980s than in the late 1970s, yet many economists have commented that real interest rates were actually much higher in the mid-1980s than in the late 1970s. Does this make sense? Do you think that these economists are right?

13. Retired persons often have much of their wealth placed in savings accounts and other interest-bearing investments and complain whenever interest rates are low. Do they have a valid complaint?

APPLIED PROBLEMS

14. If the interest rate is 10%, what is the present value of a security that pays you $1100 next year, $1210 the year after, and $1331 the year after that?

15. Calculate the present value of a $1000 discount bond with five years to maturity if the yield to maturity is 6%.

16. A lottery claims its grand prize is $10 million, payable over five years at $2 000 000 per year. If the first payment is made immediately, what is this grand prize really worth? Use an interest rate of 6%.

17. Suppose that a commercial bank wants to buy Treasury bills. These instruments pay $5000 in one year and are currently selling for $5012. What is the yield to maturity of these bonds? Is this a typical situation? Why?

18. What is the yield to maturity on a simple loan for $1 million that requires a repayment of $2 million in five years' time?

19. Which $1000 bond has the higher yield to maturity, a 20-year bond selling for $800 with a current yield of 15% or a 1-year bond selling for $800 with a current yield of 5%?

20. Consider a bond with a 4% annual coupon and a face value of $1000. Complete the following table. What relationships do you observe between years to maturity, yield to maturity, and the current price?

Years to Maturity	Yield to Maturity	Current Price
2	2%	
2	4%	
3	4%	
5	2%	
5	6%	

21. Consider a coupon bond that has a $1000 par value and a coupon rate of 10%. The bond is currently selling for $1044.89 and has two years to maturity. What is the bond's yield to maturity?

22. What is the price of a perpetuity that has a coupon of $50 per year and a yield to maturity of 2.5%? If the yield to maturity doubles, what will happen to the perpetuity's price?

23. Property taxes in a particular district are 4% of the purchase price of a home every year. If you just purchased a $250 000 home, what is the present value of all the future property tax payments? Assume that the house remains worth $250 000 forever, property tax rates never change, and a 6% interest rate is used for discounting.

24. A $1000-face-value bond has a 10% coupon rate, its current price is $960, and its price is expected to increase to $980 next year. Calculate the current yield, the expected rate of capital gain, and the expected rate of return.

25. Suppose that you want to take out a loan and that your local bank wants to charge you an annual real interest rate equal to 3%. Assuming that the annualized expected rate of inflation over the life of the bond is 1%, determine the nominal interest rate that the bank will charge you. What happens if, over the life of the loan, actual inflation is 0.5%?

DATA ANALYSIS PROBLEMS

1. Go to the CANSIM database and download monthly data from January 1953 to December 2020 for the three-month T-bill rate (series V122541) and the total consumer price index (series V41690973).

a. What is the average year-over-year inflation rate, using the formula:

$$\pi_t = 100 \times (P_t - P_{t-12})/P_{t-12}$$

b. What is the average T-bill rate?

c. Assume that the expected inflation rate is the same as the actual inflation rate (a restrictive assumption!) and calculate the average real interest rate, r.

2. Go to the CANSIM database and download monthly data (from November 1991 to January 2021) for the interest rate on long-term Canada Real Return Bonds (series V122553). This is a direct measure of the real interest rate.

a. What is the mean real interest rate over the sample period?

b. What is its standard deviation? (The standard deviation is the square root of the variance.)

3. **Real-Time Data Analysis** Go to the St. Louis Federal Reserve FRED database and find data on the interest rate on a four-year auto loan (TERMCBAUTO48NS). Assume that you borrow $20 000 to purchase a new automobile and that you finance it with a four-year loan at the most recent interest rate given in the database. If you make one payment per year for four years, what will the yearly payment be? What is the total amount that will be paid out on the $20 000 loan?

5 The Behaviour of Interest Rates

Learning Objectives

5.1 Identify the factors that affect the demand for assets.

5.2 Draw the demand and supply curves for the bond market, and identify the equilibrium interest rate.

5.3 List and describe the factors that affect the equilibrium interest rate in the bond market.

5.4 Describe the connection between the bond market and the money market through the liquidity preference framework.

5.5 List and describe the factors that affect the money market and the equilibrium interest rate.

5.6 Identify and illustrate the effects on the interest rate of changes in money growth over time.

Preview

In the early 1950s, nominal interest rates on three-month Treasury bills were about 1% at an annual rate; by 1981, they had reached over 20%. They fell below 2% in 2003, rose to 5% in 2007, and then fell to close to zero in 2008 and for many years afterward. They rose to above 2% in 2019, falling back to zero during the Covid-19 recession in 2020. What explains these substantial fluctuations in interest rates? One reason why we study money, banking, and financial markets is to provide some answers to this question.

In this chapter, we examine how the overall level of *nominal* interest rates (which we refer to simply as "interest rates") is determined and which factors influence their behaviour. We learned in Chapter 4 that interest rates are negatively related to the price of bonds, so if we can explain why bond prices change, we can also explain why interest rates fluctuate. We make use of supply and demand analysis for bond markets and markets for money to examine how interest rates change.

To derive a demand curve for assets such as money or bonds, the first step in our analysis, we must first understand what determines the demand for these assets. We do this by examining *portfolio theory*, an economic theory that outlines criteria that are important when deciding how much of an asset to buy. Armed with this theory, we can then go on to derive the demand curve for bonds or money. After deriving supply curves for these assets, we develop the concept of *market equilibrium*, which is defined as the point at which the quantity supplied equals the quantity demanded. Then we use this model to explain changes in equilibrium interest rates.

Because interest rates on different securities tend to move together, in this chapter we will proceed as if there were only one type of security and one interest rate in the entire economy. In the following chapter, we expand our analysis to look at why interest rates on different types of securities differ.

5.1 DETERMINANTS OF ASSET DEMAND

LO 5.1 Identify the factors that affect the demand for assets.

Before going on to our supply and demand analysis of the bond market and the market for money, we must first understand what determines the quantity demanded of an asset. Recall that an asset is a piece of property that is a store of value. Items

such as money, bonds, stocks, art, land, houses, farm equipment, and manufacturing machinery are all assets. Faced with the question of whether to buy and hold an asset or whether to buy one asset rather than another, an individual must consider the following factors:

1. **Wealth**, the total resources owned by the individual, including all assets
2. **Expected return** (the return expected over the next period) on one asset relative to alternative assets
3. **Risk** (the degree of uncertainty associated with the return) on one asset relative to alternative assets
4. **Liquidity** (the ease and speed with which an asset can be turned into cash) relative to alternative assets

Wealth

When we find that our wealth has increased, we have more resources available with which to purchase assets, and so, not surprisingly, the quantity of assets we demand increases. Therefore, the effect of changes in wealth on the quantity demanded of an asset can be summarized as follows: ***Holding everything else constant, an increase in wealth raises the quantity demanded of an asset.***

Expected Returns

In Chapter 4, we saw that the return on an asset (such as a bond) measures how much we gain from holding that asset. When we make a decision to buy an asset, we are influenced by what we expect the return on that asset to be. If a Bell Canada bond, for example, has a return of 15% half the time and 5% the other half, its expected return (which you can think of as the average return) is $10\% (= 0.5 \times 15\% + 0.5 \times 5\%)$. If the expected return on the Bell Canada bond rises relative to expected returns on alternative assets, then, holding everything else constant, it becomes more desirable to purchase the Bell Canada bond, and the quantity demanded increases. This can occur in either of two ways: (1) when the expected return on the Bell Canada bond rises while the return on an alternative asset—say, stock in Facebook—remains unchanged or (2) when the return on the alternative asset, the Facebook stock, falls while the return on the Bell Canada bond remains unchanged. To summarize, ***an increase in an asset's expected return relative to that of an alternative asset, holding everything else unchanged, raises the quantity demanded of the asset.***

Risk

The degree of risk or uncertainty of an asset's returns also affects demand for the asset. Consider two assets, stock in Fly-by-Night Airlines and stock in Feet-on-the-Ground Bus Company. Suppose that Fly-by-Night stock has a return of 15% half the time and 5% the other half, making its expected return 10%, while Feet-on-the-Ground stock has a fixed return of 10%. Fly-by-Night stock has uncertainty associated with its returns and so has greater risk than Feet-on-the-Ground stock, whose return is a sure thing.

A *risk-averse* person prefers stock in Feet-on-the-Ground (the sure thing) to Fly-by-Night stock (the riskier asset), even though the stocks have the same expected return, 10%. By contrast, a person who prefers risk is a *risk preferrer* or *risk lover*. Most

people are risk-averse, especially in their financial decisions: Everything else being equal, they prefer to hold the less risky asset. Hence, ***holding everything else constant, if an asset's risk rises relative to that of alternative assets, its quantity demanded will fall.***

Liquidity

Another factor that affects the demand for an asset is how quickly it can be converted into cash at low costs—its liquidity. An asset is liquid if the market in which it is traded has depth and breadth, that is, if the market has many buyers and sellers. A house is not a very liquid asset because it may be hard to find a buyer quickly; if a house must be sold to pay off bills, it might have to be sold for a much lower price. And the transaction costs associated with selling a house (broker's commissions, lawyer's fees, and so on) are substantial. A Canadian government Treasury bill, by contrast, is a highly liquid asset. It can be sold in a well-organized market with many buyers, and so it can be sold quickly at low cost. ***The more liquid an asset is relative to alternative assets, holding everything else unchanged, the more desirable it is and the greater the quantity demanded will be.***

Theory of Portfolio Choice

All the determining factors we have just discussed can be assembled into the **theory of portfolio choice**, which tells us how much of an asset people will want to hold in their portfolios. It states that, holding all other factors constant

1. The quantity demanded of an asset is positively related to wealth.
2. The quantity demanded of an asset is positively related to its expected return relative to alternative assets.
3. The quantity demanded of an asset is negatively related to the risk of its returns relative to alternative assets.
4. The quantity demanded of an asset is positively related to its liquidity relative to alternative assets.

These results are summarized in Table 5-1.

SUMMARY TABLE 5-1

Response of the Quantity of an Asset Demanded to Changes in Wealth, Expected Returns, Risk, and Liquidity

Variable	Change in Variable	Change in Quantity Demanded
Wealth	↑	↑
Expected return relative to other assets	↑	↑
Risk relative to other assets	↑	↓
Liquidity relative to other assets	↑	↑

Note: Only increases in the variables are shown. The effects of decreases in the variables on the quantity demanded would be the opposite of those indicated in the rightmost column.

5.2 SUPPLY AND DEMAND IN THE BOND MARKET

LO 5.2 Draw the demand and supply curves for the bond market, and identify the equilibrium interest rate.

Our first approach to the analysis of interest-rate determination looks at supply and demand in the bond market so that we can better understand how the prices of bonds are determined. Thanks to our knowledge from Chapter 4 of how interest rates are measured, we know that each bond price is associated with a particular level of the interest rate. Specifically, the negative relationship between bond prices and interest rates means that when a bond's price rises, its interest rate falls, and vice versa.

The first step in our analysis is to obtain a bond **demand curve**, which shows the relationship between the quantity demanded and the price when all other economic variables are held constant (that is, values of other variables are taken as given). You may recall from previous economics courses that the assumption that all other economic variables are held constant is called *ceteris paribus*, which means "other things being equal" in Latin.

Demand Curve

To clarify and simplify our analysis, let's consider the demand for one-year discount bonds, which make no coupon payments but pay the owner the $1000 face value in a year. If the holding period is one year, then, as we saw in Chapter 4, the return on the bonds is known absolutely and is equal to the interest rate as measured by the yield to maturity. This means that the expected return on this bond is equal to the interest rate i, which, using Equation 6 in Chapter 4, is

$$i = R^e = \frac{F - P}{P}$$

where
i = interest rate = yield to maturity
R^e = expected return
F = face value of the discount bond
P = initial purchase price of the discount bond

This formula shows that a particular value of the interest rate corresponds to each bond price. If the bond sells for $950, the interest rate and expected return are

$$\frac{\$1000 - \$950}{\$950} = 0.053 = 5.3\%$$

At this 5.3% interest rate and expected return corresponding to a bond price of $950, let us assume that the quantity of bonds demanded is $100 billion, which is plotted as point A in Figure 5-1.

At a price of $900, the interest rate and expected return are

$$\frac{\$1000 - \$900}{\$900} = 0.111 = 11.1\%$$

Because the expected return is higher, with all other economic variables (such as wealth, expected returns on other assets, risk, and liquidity) held constant, the quantity demanded of these bonds will be higher, as predicted by portfolio theory. Point B in Figure 5-1 shows that the quantity of bonds demanded at the price of $900 has risen to $200 billion. Continuing with this reasoning, we see that if the bond price is $850 (interest rate and expected return = 17.6%), the quantity of bonds demanded (point C)

Mini-lecture

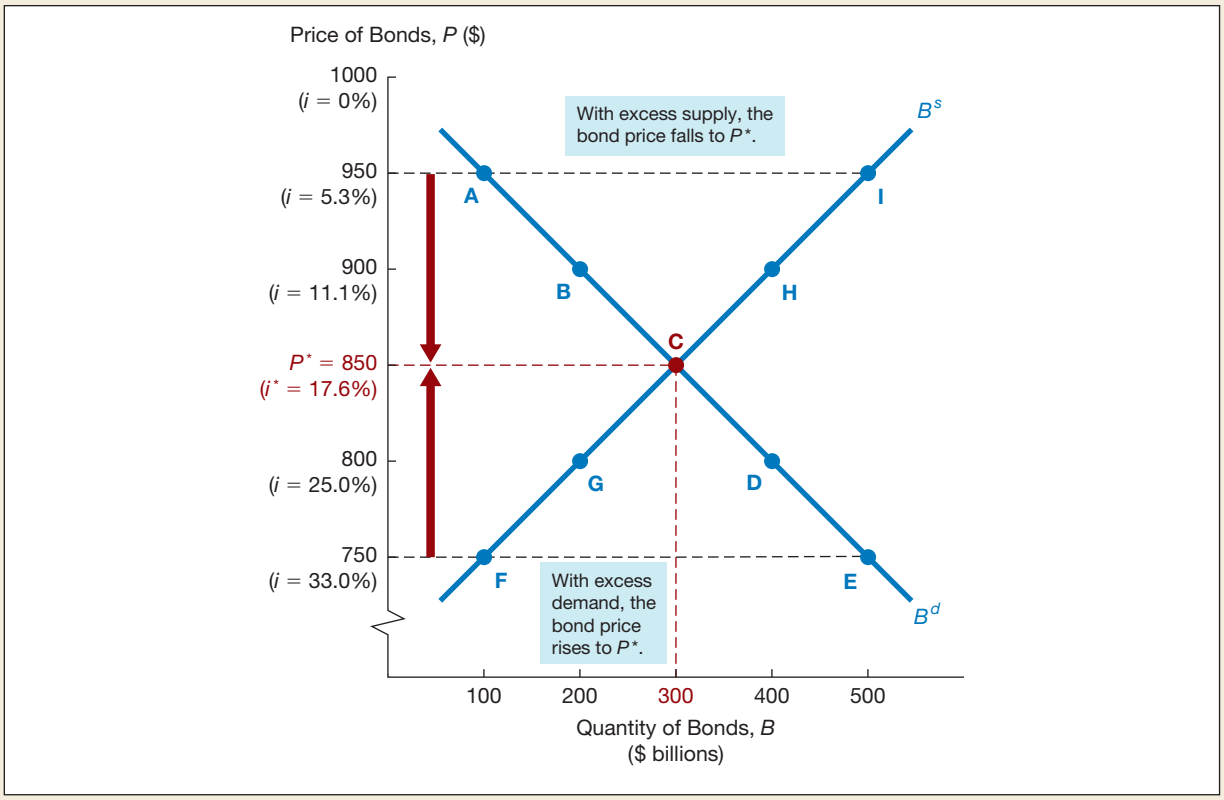


FIGURE 5-1 Supply and Demand for Bonds

Equilibrium in the bond market occurs at point C, the intersection of the demand curve B^d and the bond supply curve B^s. The equilibrium price is $P^* = \$850$, and the equilibrium interest rate is $i^* = 17.6\%$.

will be greater than at point B. Similarly, at the even lower prices of \$800 (interest rate = 25%) and \$750 (interest rate = 33.3%), the quantity of bonds demanded will be even higher (points D and E). The curve B^d, which connects these points, is the demand curve for bonds. It has the usual downward slope, indicating that at lower prices of the bond (everything else being equal), the quantity demanded is higher.[1]

Supply Curve

An important assumption behind the demand curve for bonds in Figure 5-1 is that all other economic variables besides the bond's price and interest rate are held constant. We use the same assumption in deriving a **supply curve**, which shows the relationship between the quantity supplied and the price when all other economic variables are held constant.

In Figure 5-1, when the price of the bonds is \$750 (interest rate = 33.3%), point F shows that the quantity of bonds supplied is \$100 billion for the example we are considering. If the price is \$800, the interest rate is the lower rate of 25%. Because at this interest rate it is now less costly to borrow by issuing bonds, firms will be willing to borrow

[1]Although our analysis indicates that the demand curve is downward-sloping, it does not imply that the curve is a straight line. For ease of exposition, however, we will draw demand curves and supply curves as straight lines.

more through bond issues, and the quantity of bonds supplied is at the higher level of $200 billion (point G). An even higher price of $850, corresponding to a lower interest rate of 17.6%, results in a larger quantity of bonds supplied of $300 billion (point C). Higher prices of $900 and $950 result in even lower interest rates and even greater quantities of bonds supplied (points H and I). The B^s curve, which connects these points, is the supply curve for bonds. It has the usual upward slope found in supply curves, indicating that as the price increases (everything else being equal), the quantity supplied increases.

Market Equilibrium

In economics, **market equilibrium** occurs when the amount that people are willing to buy (*demand*) equals the amount that people are willing to sell (*supply*) at a given price. In the bond market, this is achieved when the quantity of bonds demanded equals the quantity of bonds supplied:

$$B^d = B^s \tag{1}$$

In Figure 5-1, equilibrium occurs at point C, where the demand and supply curves intersect at a bond price of $850 (interest rate of 17.6%) and a quantity of bonds of $300 billion. The price of $P^* = 850$, where the quantity demanded equals the quantity supplied, is called the *equilibrium* or *market-clearing* price. Similarly, the interest rate of $i^* = 17.6\%$ that corresponds to this price is called the equilibrium or market-clearing interest rate.

The concepts of market equilibrium and equilibrium price or interest rate are useful because the market tends to head toward them. We can see this in Figure 5-1 by first looking at what happens when we have a bond price that is above the equilibrium price. When the price of bonds is set too high, at, say, $950, the quantity of bonds supplied at point I is greater than the quantity of bonds demanded at point A. A situation like this, in which the quantity of bonds supplied exceeds the quantity of bonds demanded, is called a condition of **excess supply**. Because people (borrowers) want to sell more bonds than others (lender-savers) want to buy, the price of the bonds will fall, as shown by the downward arrow in the figure at the bond price of $950. As long as the bond price remains above the equilibrium price, an excess supply of bonds will continue to be available, and the price of bonds will continue to fall. This decline will stop only when the price has reached the equilibrium price of $850, the price at which the excess supply of bonds has been eliminated.

Now let's look at what happens when the price of bonds is below the equilibrium price. If the price of the bonds is set too low, at, say, $750, the quantity demanded at point E is greater than the quantity supplied at point F. This is called a condition of **excess demand**. People (lender-savers) now want to buy more bonds than others (borrowers) are willing to sell, so the price of bonds will be driven up, as illustrated by the upward arrow in the figure at the bond price of $750. Only when the excess demand for bonds is eliminated by the bond price rising to the equilibrium level of $850 is there no further tendency for the price to rise.

We can see that the concept of equilibrium price is a useful one because it indicates where the market will settle. Because each price on the vertical axis of Figure 5-1 corresponds to a particular value of the interest rate, the same diagram also shows that the interest rate will head toward the equilibrium interest rate of 17.6%. When the interest rate is below the equilibrium interest rate, as it is when it is at 5.3%, the price of the bond is above the equilibrium price, and an excess supply of bonds results. The price of the bond then falls, leading to a rise in the interest rate toward the equilibrium level. Similarly, when the interest rate is above the equilibrium level, as it is when it is at

33.3%, an excess demand for bonds occurs, and the bond price rises, driving the interest rate back down to the equilibrium level of 17.6%.

Supply and Demand Analysis

Our Figure 5-1 is a conventional supply and demand diagram with price on the vertical axis and quantity on the horizontal axis. Because the interest rate that corresponds to each bond price is also marked on the vertical axis, this diagram allows us to read the equilibrium interest rate, giving us a model that describes the determination of interest rates. It is important to recognize that a supply and demand diagram like Figure 5-1 can be drawn for *any* type of bond because the interest rate and price of a bond are *always* negatively related for all kinds of bonds, whether a discount bond or a coupon bond.

An important feature of the analysis here is that supply and demand are always described in terms of *stocks* (amounts at a given point in time) of assets, not in terms of *flows*. The **asset market approach** for understanding behaviour in financial markets—which emphasizes stocks of assets, rather than flows, in determining asset prices—is the dominant methodology used by economists, because correctly conducting analyses in terms of flows is very tricky, especially when we encounter inflation.[2]

5.3 CHANGES IN EQUILIBRIUM INTEREST RATES

LO 5.3 List and describe the factors that affect the equilibrium interest rate in the bond market.

We will now use the supply and demand framework for bonds to analyze why interest rates change. To avoid confusion, it is important to make the distinction between *movements along* a demand (or supply) curve and *shifts in* a demand (or supply) curve. When quantity demanded (or supplied) changes as a result of a change in the price of the bond (or, equivalently, a change in the interest rate), we have a *movement along* the demand (or supply) curve. The change in the quantity demanded when we move from points A to B to C in Figure 5-1, for example, is a movement along a demand curve. A *shift in* the demand (or supply) curve, by contrast, occurs when the quantity demanded (or supplied) changes *at each given price (or interest rate)* of the bond in response to a change in some other factor besides the bond's price or interest rate. When one of these factors changes, causing a shift in the demand or supply curve, there will be a new equilibrium value for the interest rate.

In the following pages, we will look at how the supply and demand curves shift in response to changes in variables, such as expected inflation and wealth, and what effects these changes have on the equilibrium value of interest rates.

Shifts in the Demand for Bonds

The theory of portfolio choice, which we developed at the beginning of the chapter, provides a framework for deciding which factors will cause the demand curve for bonds to shift. These factors include changes in the following four parameters:

[2]The analysis of the bond market that we have developed here has another interpretation that uses a different terminology and framework involving the supply and demand for loanable funds. This loanable funds framework is discussed in the appendix to this chapter that is available on the MyLab.

1. Wealth
2. Expected returns on bonds relative to alternative assets
3. Risk of bonds relative to alternative assets
4. Liquidity of bonds relative to alternative assets

To see how a change in each of these factors (holding all other factors constant) can shift the demand curve, let's look at some examples. (As a study aid, Table 5-2 summarizes the effects of changes in these factors on the bond demand curve.)

Wealth When the economy is growing rapidly in a business cycle expansion and wealth is increasing, the quantity of bonds demanded at each bond price (or interest rate) increases, as shown in Figure 5-2. To see how this works, consider point B on the initial demand curve for bonds, B_1^d. With higher wealth, the quantity of bonds demanded at the same price must rise, to point B'. Similarly, for point D, the higher wealth causes the quantity demanded at the same bond price to rise to point D'. Continuing with this reasoning for every point on the initial demand curve B_1^d, we can see that the demand curve shifts to the right from B_1^d to B_2^d, as indicated by the arrows.

We can conclude that ***in a business cycle expansion with growing income and wealth, the demand for bonds rises and the demand curve for bonds shifts to the right.*** Applying the same reasoning, ***in a recession, when income and wealth are falling, the demand for bonds falls, and the demand curve shifts to the left.***

Another factor that affects wealth is the public's propensity to save. If households save more, wealth increases and, as we have seen, the demand for bonds rises and the

Mini-lecture

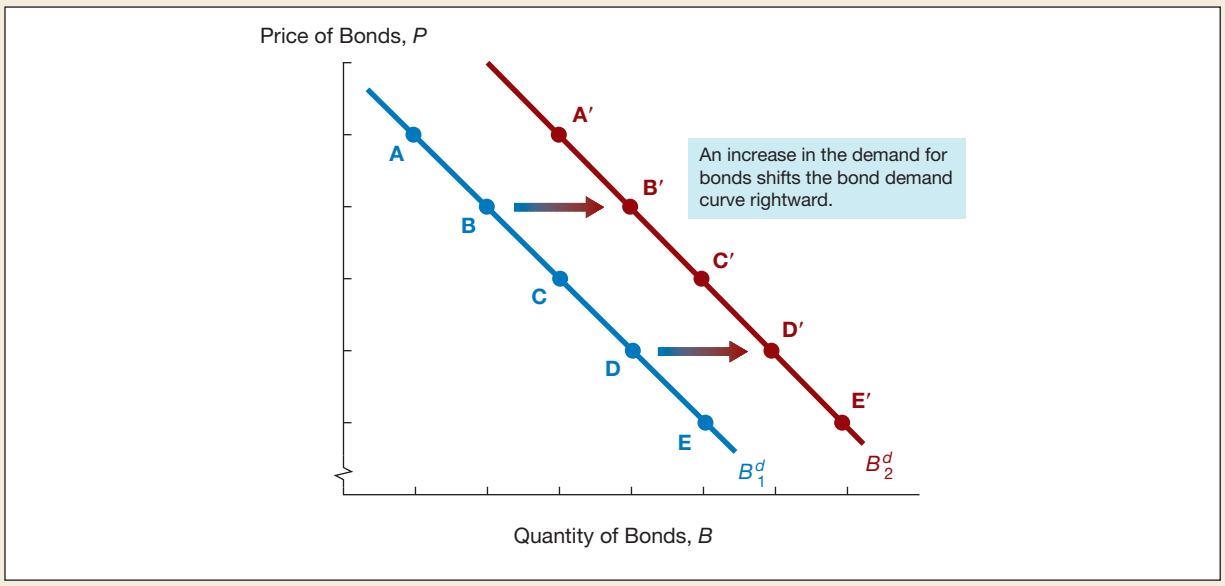
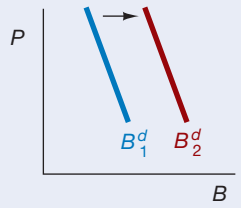
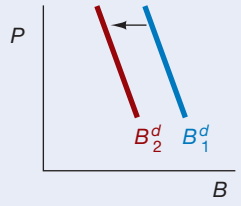
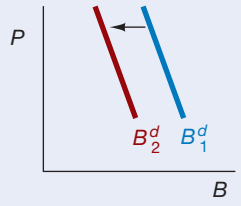
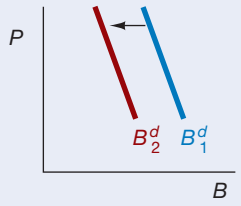
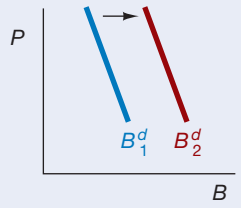


FIGURE 5-2 Shift in the Demand Curve for Bonds
When the demand for bonds increases, the demand curve shifts to the right as shown.

SUMMARY TABLE 5–2

Factors That Shift the Demand Curve for Bonds

Variable	Change in Variable	Change in Quantity Demanded at Each Bond Price	Shift in Demand Curve
Wealth	↑	↑	
Expected interest rate	↑	↓	
Expected inflation	↑	↓	
Riskiness of bonds relative to other assets	↑	↓	
Liquidity of bonds relative to other assets	↑	↑	

Note: Only increases in the variables are shown. The effects of decreases in the variables on demand would be the opposite of those indicated in the remaining columns.

demand curve for bonds shifts to the right. Conversely, if people save less, wealth and the demand for bonds fall, and the demand curve shifts to the left.

Expected Returns For a one-year discount bond and a one-year holding period, the expected return and the interest rate are identical, so nothing other than today's interest rate affects the expected return.

For bonds with maturities of greater than one year, the expected return may differ from the interest rate. For example, we saw in Chapter 4, Table 4-2, that a rise in the interest rate on a long-term bond from 10% to 20% would lead to a sharp decline in price and a very large negative return. Hence, if people began to think that interest rates would be higher next year than they had originally anticipated, the expected return today on long-term bonds would fall, and the quantity demanded would fall at each interest rate. ***Higher expected future interest rates lower the expected return for long-term bonds, decrease the demand, and shift the demand curve to the left.***

By contrast, an expectation of lower future interest rates would mean that long-term bond prices would be expected to rise more than originally anticipated, and the resulting higher expected return today would raise the quantity demanded at each bond price and interest rate. ***Lower expected future interest rates increase the demand for long-term bonds and shift the demand curve to the right*** (as in Figure 5-2).

Changes in expected returns on other assets can also shift the demand curve for bonds. If people suddenly become more optimistic about the stock market and begin to expect higher stock prices in the future, both expected capital gains and expected returns on stocks will rise. With the expected return on bonds held constant, the expected return on bonds today relative to stocks will fall, lowering the demand for bonds and shifting the demand curve to the left. ***An increase in expected return on alternative assets lowers the demand for bonds and shifts the demand curve to the left.***

A change in expected inflation is likely to alter expected returns on physical assets (also called *real assets*), such as automobiles and houses, which affect the demand for bonds. An increase in expected inflation from, say, 5% to 10% will lead to higher prices on cars and houses in the future and hence higher nominal capital gains. The resulting rise in the expected returns today on these real assets will lead to a fall in the expected return on bonds relative to the expected return on real assets today and thus cause the demand for bonds to fall. Alternatively, we can think of the rise in expected inflation as lowering the real interest rate on bonds, and thus the resulting decline in the relative expected return on bonds will cause the demand for bonds to fall. ***An increase in the expected rate of inflation lowers the expected return on bonds, causing their demand to decline and the demand curve to shift to the left.***

Risk If prices in the bond market become more volatile, the risk associated with bonds increases, and bonds become a less attractive asset. ***An increase in the riskiness of bonds causes the demand for bonds to fall and the demand curve to shift to the left.***

Conversely, an increase in the volatility of prices in another asset market, such as the stock market, would make bonds more attractive. ***An increase in the riskiness of alternative assets causes the demand for bonds to rise and the demand curve to shift to the right*** (as in Figure 5-2).

Liquidity If more people started trading in the bond market, and as a result it became easier to sell bonds quickly, the increase in their liquidity would cause the quantity of bonds demanded at each interest rate to rise. ***Increased liquidity of bonds results in an increased demand for bonds, and the demand curve shifts to the right*** (see Figure 5-2). ***Similarly, increased liquidity of alternative assets lowers the demand***

for bonds and shifts the demand curve to the left. The reduction of brokerage commissions for trading common stocks that occurred when the fixed-rate commission structure was abolished in 1975, for example, increased the liquidity of stocks relative to bonds, and the resulting lower demand for bonds shifted the demand curve to the left.

Shifts in the Supply of Bonds

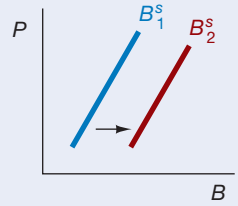
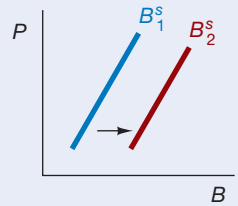
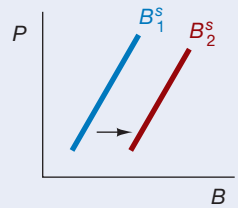
Certain factors can cause the supply curve for bonds to shift. Among these factors are the following:

1. Expected profitability of investment opportunities
2. Expected inflation
3. Government budget deficits

We will look at how the supply curve shifts when each of these factors changes (all others remaining constant). (As a study aid, Table 5-3 summarizes the effects of changes in these factors on the bond supply curve.)

SUMMARY TABLE 5-3

Factors That Shift the Supply Curve for Bonds

Variable	Change in Variable	Change in Quantity Supplied at Each Bond Price	Shift in Supply Curve
Profitability of investments	↑	↑	
Expected inflation	↑	↑	
Government deficit	↑	↑	

Note: Only increases in the variables are shown. The effects of decreases in the variables on the supply would be the opposite of those indicated in the remaining columns.

Mini-lecture

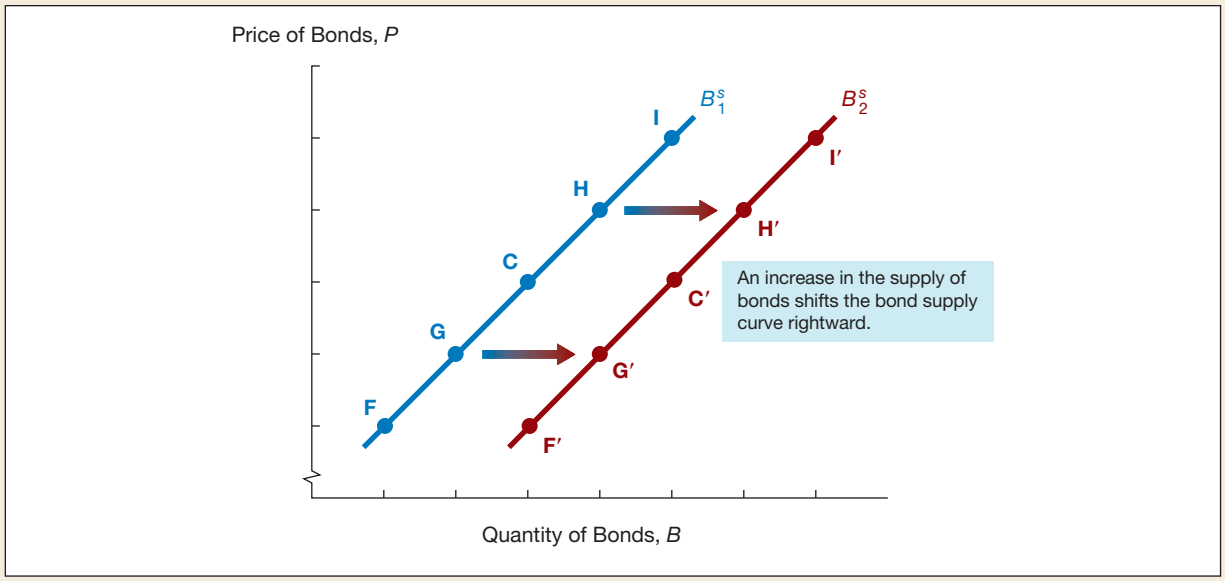


FIGURE 5-3 Shift in the Supply Curve for Bonds
When the supply of bonds increases, the supply curve shifts to the right.

Expected Profitability of Investment Opportunities When opportunities for profitable plant and equipment investments are plentiful, firms are more willing to borrow to finance these investments. When the economy is growing rapidly, as in a business cycle expansion, investment opportunities that are expected to be profitable abound, and the quantity of bonds supplied at any given bond price increases (for example, from G to G′ or H to H′ in Figure 5-3). ***Therefore, in a business cycle expansion, the supply of bonds increases and the supply curve shifts to the right. Likewise, in a recession, when far fewer profitable investment opportunities are expected, the supply of bonds falls and the supply curve shifts to the left.***

Expected Inflation As we saw in Chapter 4, the real cost of borrowing is most accurately measured by the real interest rate, which equals the (nominal) interest rate minus the expected inflation rate. For a given interest rate (and bond price), when expected inflation increases, the real cost of borrowing falls; hence, the quantity of bonds supplied increases at any given bond price. ***An increase in expected inflation causes the supply of bonds to increase and the supply curve to shift to the right*** (see Figure 5-3), ***and a decrease in expected inflation causes the supply of bonds to decrease and the supply curve to shift to the left.***

Government Budget Deficits The activities of the government can influence the supply of bonds in several ways. The Canadian government Treasury issues bonds to finance government deficits, caused by gaps between the government's expenditures and its revenues. When these deficits are large, the government sells more bonds, and the quantity of bonds supplied at each bond price increases. ***Higher government deficits increase the supply of bonds and shift the supply curve to the right*** (see Figure 5-3). ***On the other hand, government surpluses, as occurred in the late 1990s, decrease the supply of bonds and shift the supply curve to the left.***

Provincial and municipal governments and other government agencies also issue bonds to finance their expenditures, and this can affect the supply of bonds as well. We will see in later chapters that the conduct of monetary policy involves the purchase and sale of bonds, which in turn influences the supply of bonds.

We now can use our knowledge of how supply and demand curves shift to analyze how the equilibrium interest rate can change. The best way to do this is to pursue several applications that are particularly relevant to our understanding of how monetary policy affects interest rates. In studying these applications, keep two things in mind:

1. When we examine the effect of a variable change, remember we are assuming that all other variables are unchanged; that is, we are making use of the *ceteris paribus* assumption.
2. Remember that the interest rate is negatively related to the bond price, so when the equilibrium bond price rises, the equilibrium interest rate falls. Conversely, if the equilibrium bond price moves downward, the equilibrium interest rate rises.

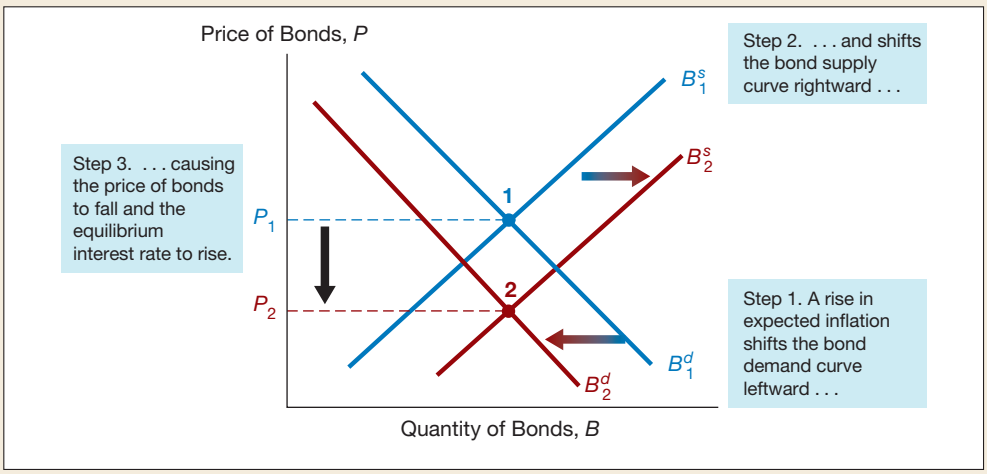
APPLICATION Changes in the Interest Rate Due to a Change in Expected Inflation: The Fisher Effect

We have already done most of the work necessary to evaluate how a change in expected inflation affects the nominal interest rate, in that we have already analyzed how a change in expected inflation shifts the supply and demand curves for bonds. Figure 5-4 shows the effect on the equilibrium interest rate of an increase in expected inflation.

Mini-lecture

FIGURE 5-4 Response to a Change in Expected Inflation

When expected inflation rises, the supply curve shifts from B_1^s to B_2^s, and the demand curve shifts from B_1^d to B_2^d. The equilibrium moves from point 1 to point 2, causing the equilibrium bond price to fall from P_1 to P_2 and the equilibrium interest rate to rise.


Price of Bonds, P
Step 2. . . . and shifts the bond supply curve rightward . . .
Step 3. . . . causing the price of bonds to fall and the equilibrium interest rate to rise.
B_1^s
B_2^s
P_1
1
P_2
2
Step 1. A rise in expected inflation shifts the bond demand curve leftward . . .
B_1^d
B_2^d
Quantity of Bonds, B

Suppose that expected inflation is initially 5% and the initial supply and demand curves B_1^s and B_1^d intersect at point 1, where the equilibrium bond price is P_1. If expected inflation rises to 10%, the expected return on bonds relative to real assets falls for any given bond price and interest rate. As a result, the demand for bonds falls, and the demand curve shifts to the left from B_1^d to B_2^d. The rise in expected inflation also shifts the supply curve. At any given bond price and interest rate, the real cost of borrowing declines, causing the quantity of bonds supplied to increase and the supply curve to shift to the right, from B_1^s to B_2^s.

When the demand and supply curves shift in response to the rise in expected inflation, the equilibrium moves from point 1 to point 2, at the intersection of B_2^d and B_2^s. The equilibrium bond price falls from P_1 to P_2 and, because the bond price is negatively related to the interest rate, this means that the equilibrium interest rate rises. Note that Figure 5-4 has been drawn so that the equilibrium quantity of bonds remains the same at both point 1 and point 2. However, depending on the size of the shifts in the supply and demand curves, the equilibrium quantity of bonds can either rise or fall when expected inflation rises.

Our supply and demand analysis has led us to an important observation: ***When expected inflation rises, interest rates will rise.*** This result has been named the **Fisher effect**, after Irving Fisher, the economist who first pointed out the relationship of expected inflation to interest rates. The accuracy of this prediction is shown in Figure 5-5. The interest rate on three-month Treasury bills has usually moved along with the expected inflation rate. Consequently, many economists recommend that inflation must be kept low if we want to keep nominal interest rates low. ◆

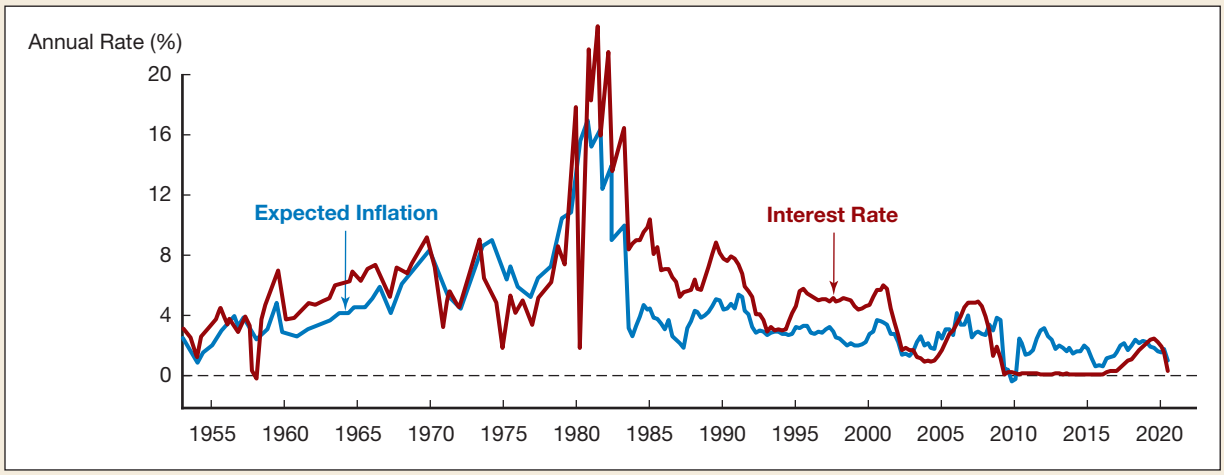


FIGURE 5-5 Expected Inflation and Interest Rates (Three-Month Treasury Bills), 1953–2020

The interest rate on three-month Treasury bills and the expected inflation rate generally move together, as the Fisher effect predicts.

Sources: Federal Reserve Bank of St. Louis FRED database: fred.stlouisfed.org/series/TB3MS; https://fred.stlouisfed.org/series/CPIAUCSL/

Expected inflation calculated using procedures outlined in Frederic S. Mishkin, "The Real Interest Rate: An Empirical Investigation," *Carnegie-Rochester Conference Series on Public Policy* 15 (1981): 151–200. These procedures involve estimating expected inflation as a function of past interest rates, inflation, and time trends.

APPLICATION Changes in the Interest Rate Due to a Business Cycle Expansion

Figure 5-6 analyzes the effects of a business cycle expansion on interest rates. In a business cycle expansion, the amounts of goods and services being produced in the economy increase, so national income rises. When this occurs, businesses are more willing to borrow because they are likely to have many profitable investment opportunities for which they need financing. Hence, at a given bond price, the quantity of bonds that firms want to sell (that is, the supply of bonds) will increase. This means that during a business cycle expansion, the supply curve for bonds shifts to the right (see Figure 5-6) from B_1^s to B_2^s.

Expansion in the economy also affects the demand for bonds. As the economy expands, wealth is likely to increase, and the theory of portfolio choice tells us that the demand for bonds will rise as well. We see this in Figure 5-6, where the demand curve has shifted to the right, from B_1^d to B_2^d.

Given that both the supply and demand curves have shifted to the right, we know that the new equilibrium reached at the intersection of B_2^d and B_2^s must also move to the right. However, depending on whether the supply curve shifts more than the demand curve or vice versa, the new equilibrium interest rate can either rise or fall.

The supply and demand analysis used here gives us an ambiguous answer to the question of what happens to interest rates in a business cycle expansion. Figure 5-6 has been drawn so that the shift in the supply curve is greater than the shift in the demand curve,

Mini-lecture

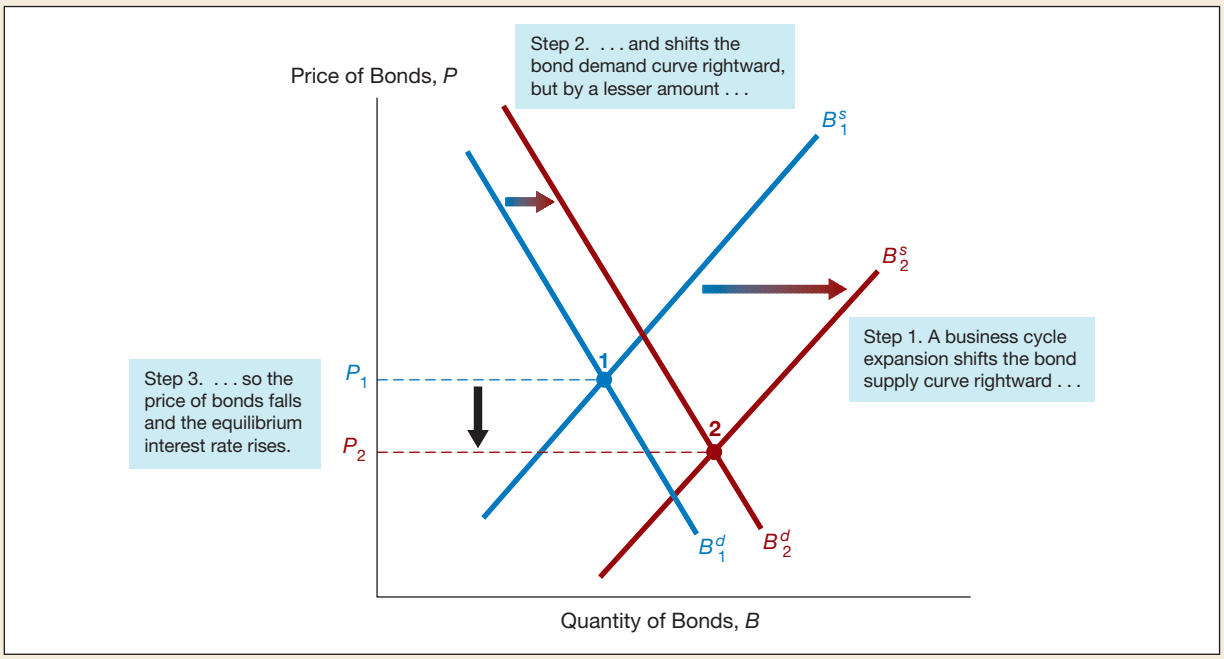

Step 2. . . . and shifts the bond demand curve rightward, but by a lesser amount . . .
Price of Bonds, P
B_1^s
B_2^s
Step 1. A business cycle expansion shifts the bond supply curve rightward . . .
Step 3. . . . so the price of bonds falls and the equilibrium interest rate rises.
P_1
P_2
1
2
B_1^d
B_2^d
Quantity of Bonds, B

FIGURE 5-6 Response to a Business Cycle Expansion

In a business cycle expansion, when income and wealth are rising, the demand curve shifts rightward from B_1^d to B_2^d. If the supply curve shifts to the right more than the demand curve, as in this figure, the equilibrium bond price moves down from P_1 to P_2 and the equilibrium interest rate rises.

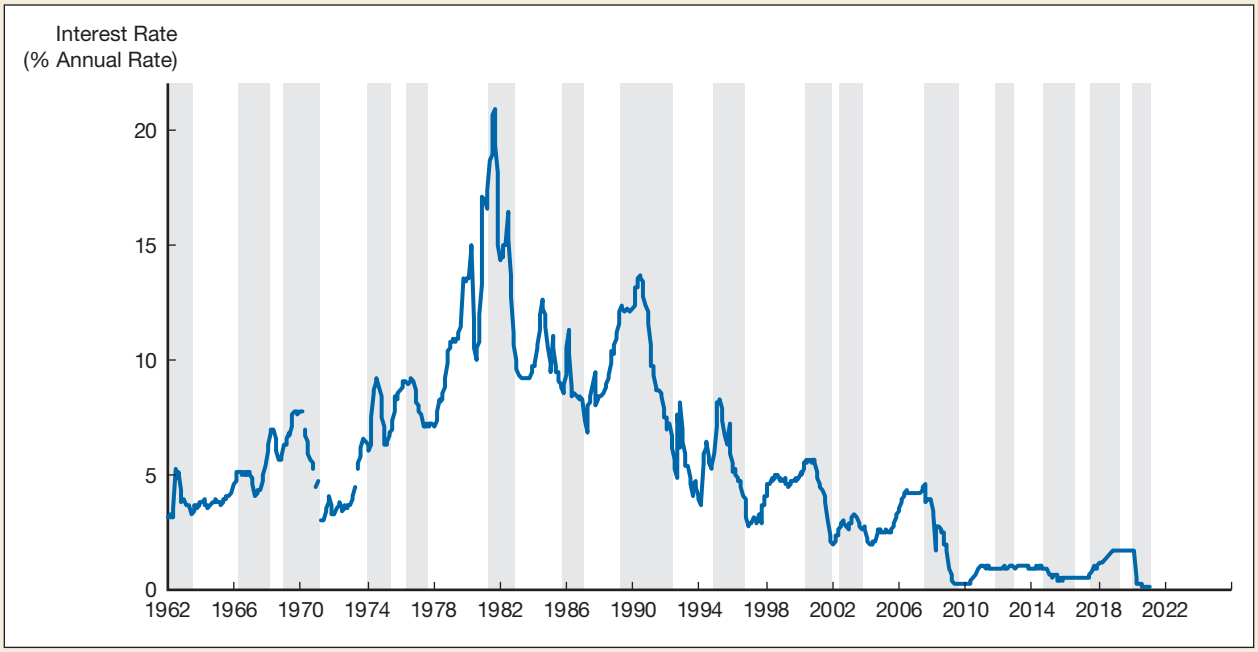


FIGURE 5-7 Business Cycles and Interest Rates (Three-Month Treasury Bills), 1962–2020
Shaded areas indicate periods of recession. The interest rate tends to rise during business cycle expansions and fall during recessions.
Source: Statistics Canada CANSIM series V1225312.

causing the equilibrium bond price to fall to P_2 and the equilibrium interest rate to rise. The reason the figure has been drawn such that a business cycle expansion and a rise in income lead to a higher interest rate is that this is the outcome we actually see in the data. Figure 5-7 plots the movement of the interest rate on three-month Canadian Treasury bills from 1962 to 2020 and indicates when the business cycle went through recessions (shaded areas). As you can see, the interest rate tends to rise during business cycle expansions and fall during recessions, as indicated by the supply and demand diagram in Figure 5-6. ◆

APPLICATION Explaining Current Low Interest Rates in Europe, Japan, and the United States: Low Inflation and Secular Stagnation

In the aftermath of the global financial crisis, interest rates in Canada, Europe, and the United States, as well as in Japan, have fallen to extremely low levels. They fell even lower during the Covid-19 recession in 2020. Indeed, as discussed in Chapter 4, we have seen that interest rates have even sometimes turned negative. Why are interest rates in these countries at such low levels?

The answer is that inflation has fallen to very low levels in all these countries, sometimes even going negative, while at the same time there has been a dearth of attractive investment opportunities. Using these facts, analysis similar to that used in the preceding applications explains why interest rates have become so low.

Very low and even negative inflation causes the demand for bonds to rise because the expected return on real assets falls, thereby raising the relative expected return on bonds and in turn causing the demand curve to shift to the right. Low and even negative inflation also raises the real interest rate and therefore the real cost of borrowing for any given nominal rate, thereby causing the supply of bonds to contract and the supply curve to shift to the left. The outcome is then exactly the opposite of that graphed in Figure 5-4: The rightward shift of the demand curve and leftward shift of the supply curve lead to a rise in the bond price and a fall in interest rates.

All of these countries have also been experiencing very low economic growth rates and a lack of profitable investment opportunities. The former president of Harvard University and former U.S. Treasury Secretary, Lawrence Summers, has described this phenomenon with the phrase "secular stagnation." As a result of secular stagnation, firms have cut back on their investment spending, with the result that they supply fewer bonds, shifting the supply curve to the left. The leftward shift of the supply curve for bonds leads to a further rise in the bond price and interest rates fall (the opposite outcome to that in Figure 5-6).

Usually we think that low interest rates are a good thing because they make it cheap to borrow. But the recent episodes of low interest rates in the United States, Europe, and Japan show that just as a fallacy is present in the adage "You can never be too rich or too thin" (maybe you can't be too rich, but you can certainly be too thin and thereby damage your health), a fallacy is present in always thinking that lower interest rates are better. In the United States, Europe, and Japan, the low and even negative interest rates are a sign that these economies are not doing all that well, with slow growth and inflation that is too low. ◆

5.4 SUPPLY AND DEMAND IN THE MARKET FOR MONEY: THE LIQUIDITY PREFERENCE FRAMEWORK

LO 5.4 Describe the connection between the bond market and the money market through the liquidity preference framework.

An alternative model for determining the equilibrium interest rate, developed by John Maynard Keynes, is known as the **liquidity preference framework**. This framework determines the equilibrium interest rate in terms of the supply of and demand for money rather than the supply of and demand for bonds. Although the two frameworks look different, the liquidity preference analysis of the market for money is closely related to the supply and demand framework of the bond market.[3]

The starting point of Keynes's analysis is his assumption that people use two main categories of assets to store their wealth: money and bonds. Therefore, total wealth in the economy must equal the total quantity of bonds plus money in the economy,

[3]Note that the term *market for money* refers to the market for the medium of exchange, money. This market differs from the *money market* referred to by finance practitioners, which, as discussed in Chapter 2, is the financial market in which short-term debt instruments are traded.

which equals the quantity of bonds supplied (B^s) plus the quantity of money supplied (M^s). The quantity of bonds demanded (B^d) plus the quantity of money demanded (M^d) must also equal the total amount of wealth, because people cannot purchase more assets than their available resources allow. Thus the quantity of bonds and money supplied must equal the quantity of bonds and money demanded:

$$B^s + M^s = B^d + M^d \tag{2}$$

Collecting the bond terms on one side of the equation and the money terms on the other, this equation can be rewritten as

$$B^s - B^d = M^d - M^s \tag{3}$$

Equation 3 tells us that if the market for money is in equilibrium $(M^s = M^d)$, then the right-hand side of the equation equals zero, implying that $B^s = B^d$, which in turn means the bond market is also in equilibrium.

Thus we arrive at the same result whether we determine the equilibrium interest rate by equating the supply and demand for bonds or by equating the supply and demand for money. In this sense, the liquidity preference framework, which analyzes the market for money, is equivalent to a framework analyzing supply and demand in the bond market. In practice, the approaches differ because, by assuming there are only two kinds of assets, money and bonds, the liquidity preference approach implicitly ignores any effects on interest rates that arise from changes in the expected returns on real assets such as automobiles and houses. In most instances, however, both frameworks yield the same predictions.

The reason we approach the determination of interest rates using both frameworks is that the bond supply and demand framework is easier to use when analyzing the effects caused by changes in expected inflation, whereas the liquidity preference framework is easier to use when analyzing the effects caused by changes in income, the price level, and the supply of money.

Because the definition of money used by Keynes includes currency (which earns no interest) and chequing account deposits (which in his time typically earned little or no interest), he assumed that money has a zero rate of return. Bonds, the only alternative asset to money in Keynes's framework, have an expected return equal to the interest rate i.[4] As this interest rate rises (holding everything else unchanged), the expected return on money falls relative to the expected return on bonds, causing a fall in the quantity of money demanded, as predicted by the theory of portfolio choice.

We can also see that the quantity of money demanded and the interest rate should be negatively related by using the concept of **opportunity cost**, which is the amount of interest (expected return) sacrificed by not holding the alternative asset—in this case, a bond. As the interest rate on bonds, i, rises, the opportunity cost of holding money rises; thus money is less desirable and the quantity of money demanded falls.

Figure 5-8 shows the quantity of money demanded at a number of interest rates, with all other economic variables, such as income and the price level, held constant. At an interest rate of 25%, point A shows that the quantity of money demanded is $100 billion. If the interest rate is at the lower rate of 20%, the opportunity cost of

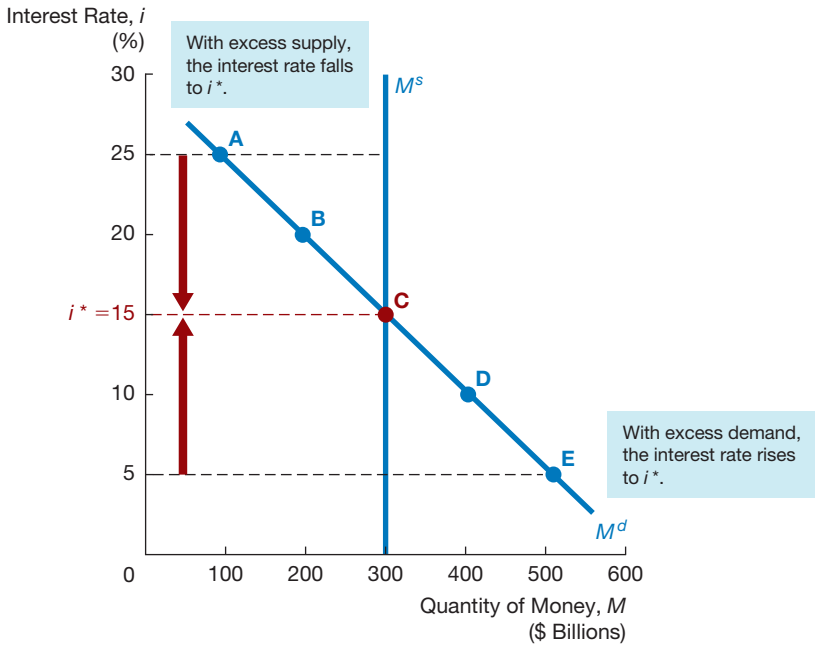
[4]Keynes did not actually assume that the expected returns on bonds equalled the interest rate but rather argued that they were closely related. This distinction makes no appreciable difference in our analysis.

Mini-lecture

FIGURE 5-8

Equilibrium in the Market for Money

Equilibrium in the market for money occurs at point C, the intersection of the money demand curve M^d and the money supply curve M^s. The equilibrium interest rate is $i^* = 15\%$.



holding money is lower, and the quantity of money demanded rises to $200 billion, as indicated by the move from point A to point B. If the interest rate is even lower, the quantity of money demanded is even higher, as is indicated by points C, D, and E. The curve M^d connecting these points is the demand curve for money, and it slopes downward.

At this point in our analysis, we will assume that a central bank controls the amount of money supplied at a fixed quantity of $300 billion, so the supply curve for money M^s in the figure is a vertical line at $300 billion. The equilibrium at which the quantity of money demanded equals the quantity of money supplied occurs at the intersection of the supply and demand curves at point C, where

$$M^d = M^s \tag{4}$$

The resulting equilibrium interest rate is $i^* = 15\%$.

We can again see that there is a tendency to approach this equilibrium rate by first looking at the relationship of money demand and supply when the interest rate is above the equilibrium interest rate. When the interest rate is 25%, the quantity of money demanded at point A is $100 billion, yet the quantity of money supplied is $300 billion. The excess supply of money means that people are holding more money than they desire, so they will try to get rid of their excess money balances by trying to buy bonds. Accordingly, they will bid up the price of bonds. As the bond price rises, the interest rate will fall toward the equilibrium interest rate of 15%. This tendency is shown by the downward arrow drawn at the interest rate of 25%.

Likewise, if the interest rate is 5%, the quantity of money demanded at point E is $500 billion, but the quantity of money supplied is only $300 billion. An excess demand for money now exists because people want to hold more money than they currently have. To try to obtain more money, they will sell their only other asset—bonds—and the price will fall. As the price of bonds falls, the interest rate will rise toward the equilibrium rate of 15%. Only when the interest rate is at its equilibrium value will there be no tendency for it to move further, and so the interest rate will settle at this value.

5.5 CHANGES IN EQUILIBRIUM INTEREST RATES IN THE LIQUIDITY PREFERENCE FRAMEWORK

LO 5.5 List and describe the factors that affect the money market and the equilibrium interest rate.

To use the liquidity preference framework to analyze how the equilibrium interest rate changes, we must understand what causes the demand and supply curves for money to shift.

Shifts in the Demand for Money

In Keynes's liquidity preference analysis, two factors cause the demand curve for money to shift: income and the price level.

Income Effect In Keynes's view, there were two reasons why income would affect the demand for money. First, as an economy expands and income rises, wealth increases and people want to hold more money as a store of value. Second, as the economy expands and income rises, people want to carry out more transactions using money as a medium of exchange, and so they also want to hold more money. The conclusion is that ***a higher level of income causes the demand for money at each interest rate to increase and the demand curve to shift to the right.***

Price-Level Effect Keynes took the view that people care about the amount of money they hold in real terms—that is, in terms of the goods and services it can buy. When the price level rises, the same nominal quantity of money is no longer as valuable; it cannot be used to purchase as many real goods or services. To restore their holdings of money in real terms to the former level, people will want to hold a greater nominal quantity of money, so ***a rise in the price level causes the demand for money at each interest rate to increase and the demand curve to shift to the right.***

Shifts in the Supply of Money

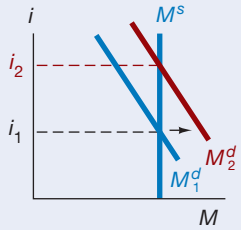
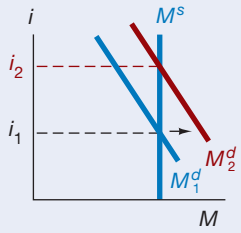
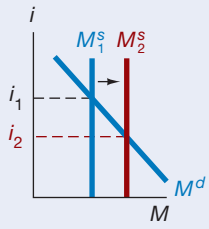
We will assume that the supply of money is completely controlled by the central bank, which in Canada is the Bank of Canada. (Actually, the process that determines the money supply is substantially more complicated, involving banks, depositors, and borrowers from banks. We will study it in more detail later in the book.) For now, all we need to know is that ***an increase in the money supply engineered by the Bank of Canada will shift the supply curve for money to the right.***

APPLICATION Changes in the Equilibrium Interest Rate Due to Changes in Income, the Price Level, or the Money Supply

To see how we can use the liquidity preference framework to analyze the movement of interest rates, we will again look at several applications that will help us evaluate the effect of monetary policy on interest rates. In going through these applications, remember to use the *ceteris paribus* assumption: When examining the effect of a change in one variable, hold all other variables constant. (As a study aid, Table 5-4 summarizes the shifts in the demand and supply curves for money.)

SUMMARY TABLE 5-4

Factors That Shift the Demand for and Supply of Money

Variable	Change in Variable	Change in Money Demand (M^d) or Supply (M^s) at Each Interest Rate	Change in Interest Rate	
Income	↑	M^d↑	↑	
Price level	↑	M^d↑	↑	
Money supply	↑	M^s↑	↓	

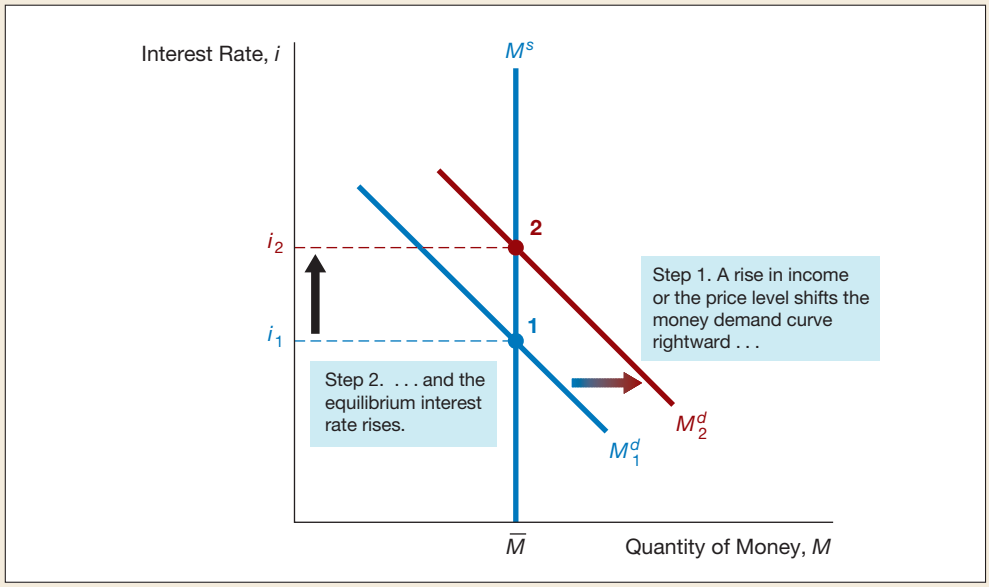
Note: Only increases in the variables are shown. The effects of decreases in the variables on demand and supply would be the opposite of those indicated in the remaining columns.

Mini-lecture

FIGURE 5-9

Response to a Change in Income or the Price Level

In a business cycle expansion, when income is rising, or when the price level rises, the demand curve shifts from M_1^d to M_2^d. The supply curve is fixed at $M^s = \overline{M}$. The equilibrium interest rate rises from i_1 to i_2.



Changes in Income

We have seen that when income is rising during a business cycle expansion, the demand for money will rise, shown in Figure 5-9 by the rightward shift in the demand curve from M_1^d to M_2^d. The new equilibrium is reached at point 2, at the intersection of the M_2^d curve with the money supply curve M^s. As you can see, the equilibrium interest rate rises from i_1 to i_2. The liquidity preference framework thus generates the conclusion that ***when income is rising during a business cycle expansion (holding other economic variables constant), interest rates will rise.*** This conclusion is unambiguous, unlike the conclusion we reached using the bond demand and supply framework.

Changes in the Price Level

When the price level rises, the value of money in terms of its purchasing power is lower. To restore their money's purchasing power in real terms to its former level, people will want to hold a greater nominal quantity of money. A higher price level shifts the demand curve for money to the right from M_1^d to M_2^d (see Figure 5-9). The equilibrium moves from point 1 to point 2, where the equilibrium interest rate has risen from i_1 to i_2, illustrating that ***when the price level increases, with the supply of money and other economic variables held constant, interest rates will rise.***

Changes in the Money Supply

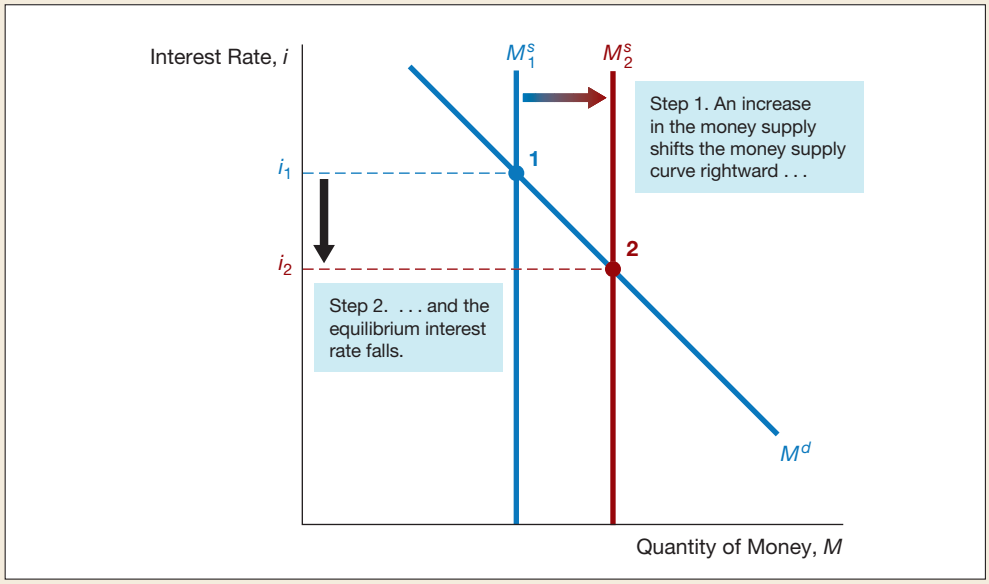
An increase in the money supply due to expansionary monetary policy by the Bank of Canada implies that the supply curve for money will shift to the right. As shown in Figure 5-10 by the movement of the supply curve from M_1^s to M_2^s, the equilibrium moves

Mini-lecture

FIGURE 5-10

Response to a Change in the Money Supply

When the money supply increases, the supply curve shifts from M_1^s to M_2^s and the equilibrium interest rate falls from i_1 to i_2.



from point 1 down to point 2, where the M_2^s supply curve intersects with the demand curve M^d and the equilibrium interest rate has fallen from i_1 to i_2. ***When the money supply increases (everything else remaining equal), interest rates will decline.***[5] ◆

5.6 MONEY AND INTEREST RATES

LO 5.6 Identify and illustrate the effects on the interest rate of changes in money growth over time.

The liquidity preference analysis in Figure 5-10 seems to suggest that an increase in the money supply will lower interest rates. This conclusion has important policy implications because it has frequently caused politicians to call for a more rapid growth of the money supply in an effort to drive down interest rates.

But is it correct to conclude that money and interest rates should be negatively related? Might other important factors have been left out of the liquidity preference analysis in Figure 5-10, factors that would reverse this conclusion? We will provide answers to these questions by applying the supply and demand analysis we used in this chapter to obtain a deeper understanding of the relationship between money and interest rates.

An important criticism of the idea that an increase in the money supply lowers interest rates was raised by Milton Friedman, a Nobel laureate in economics. He acknowledged that the liquidity preference analysis was correct and called the

[5]This same result can be generated using the bond supply and demand framework. As we will see in Chapter 15, a central bank increases the money supply primarily by buying bonds and thereby decreasing the supply of bonds available to the public. The resulting shift to the left of the supply curve for bonds leads to an increase in the equilibrium price of bonds and a decline in the equilibrium interest rate.

result—that an increase in the money supply (*everything else remaining equal*) lowers interest rates—the *liquidity effect*. However, he viewed the liquidity effect as merely part of the story: An increase in the money supply might not leave "everything else equal" and will have other effects on the economy that may make interest rates rise. If these effects are substantial, it is entirely possible that when the money supply increases, interest rates might also increase.

We have already laid the groundwork for our discussion of these other effects, because we have shown how changes in income, the price level, and expected inflation influence the equilibrium interest rate.

1. *Income Effect.* An increasing money supply can cause national income and wealth to rise. Both the liquidity preference and bond supply and demand frameworks indicate that interest rates will then rise (see Figures 5-6 and 5-9). Thus ***the income effect of an increase in the money supply is a rise in interest rates in response to the higher level of income.***
2. *Price-Level Effect.* An increase in the money supply can also cause the overall price level in the economy to rise. The liquidity preference framework predicts that this will lead to a rise in interest rates. Thus ***the price-level effect from an increase in the money supply is a rise in interest rates in response to the rise in price level.***
3. *Expected-Inflation Effect.* The higher inflation rate that can result from an increase in the money supply can also affect interest rates by influencing the expected inflation rate. Specifically, an increase in the money supply may lead people to expect a higher price level in the future—and hence the expected inflation rate will be higher. The bond supply and demand framework has shown us that this increase in expected inflation will lead to a higher level of interest rates. Therefore, ***the expected-inflation effect of an increase in the money supply is a rise in interest rates in response to the rise in the expected inflation rate.***

At first glance it might appear that the price-level effect and the expected-inflation effect are the same thing. They both indicate that increases in the price level induced by an increase in the money supply will raise interest rates. However, there is a subtle difference between the two, and this is why they are discussed as two separate effects.

Suppose that a one-time increase in the money supply today leads to a rise in prices to a permanently higher level by next year. As the price level rises over the course of this year, the interest rate will rise via the price-level effect. Only at the end of the year, when the price level has risen to its peak, will the price-level effect be at a maximum.

The rising price level will also raise interest rates via the expected-inflation effect, because people will expect that inflation will be higher over the course of the year. However, when the price level stops rising next year, inflation and the expected inflation rate will return to zero. Any rise in interest rates as a result of the earlier rise in expected inflation will then be reversed. We thus see that in contrast to the price-level effect, which reaches its greatest impact next year, the expected-inflation effect will have its smallest impact (zero impact) next year. The basic difference between the two effects, then, is that the price-level effect remains even after prices have stopped rising, whereas the expected-inflation effect disappears.

An important point is that the expected-inflation effect will persist only as long as the price level continues to rise. As we will see in our discussion of monetary theory in subsequent chapters, a one-time increase in the money supply will not produce a continually rising price level; only a higher rate of money supply growth will. Thus a higher rate of money supply growth is needed if the expected-inflation effect is to persist.

APPLICATION Does a Higher Rate of Growth of the Money Supply Lower Interest Rates?

We can now put together all the effects we have discussed to help us decide whether our analysis supports the politicians who advocate a greater rate of growth of the money supply when they feel that interest rates are too high. Of all the effects, only the liquidity effect indicates that a higher rate of money growth will cause a decline in interest rates. In contrast, the income, price-level, and expected-inflation effects indicate that interest rates will rise when money growth is higher. Which of these effects is largest, and how quickly does each take effect? The answers are critical in determining whether interest rates will rise or fall when money supply growth is increased.

Generally, the liquidity effect from greater money growth takes effect immediately, because the rising money supply leads to an immediate decline in the equilibrium interest rate. The income and price-level effects take longer to work because time is needed for the increasing money supply to raise the price level and income, which in turn raise interest rates. The expected-inflation effect, which also raises interest rates, can be slow or fast, depending on whether people adjust their expectations of inflation slowly or quickly when the money growth rate is increased.

Three possibilities are outlined in Figure 5-11; each shows how interest rates respond over time to an increased rate of money supply growth, starting at time T. Panel (a) shows a case in which the liquidity effect dominates the other effects so that the interest rate falls from i_1 at time T to a final level of i_2. The liquidity effect operates quickly to lower the interest rate, but as time goes by, the other effects start to reverse some of the decline. Because the liquidity effect is larger than the others, however, the interest rate never rises back to its initial level.

In panel (b), the liquidity effect is smaller than the other effects, with the expected-inflation effect operating slowly because expectations of inflation are slow to adjust upward. Initially, the liquidity effect drives down the interest rate. Then the income, price-level, and expected-inflation effects begin to raise it. Because these effects are dominant, the interest rate eventually rises above its initial level to i_2. In the short run, lower interest rates result from increased money growth, but eventually they end up climbing above the initial level.

In panel (c), the expected-inflation effect dominates and operates rapidly because people quickly raise their expectations of inflation when the rate of money growth increases. The expected-inflation effect begins immediately to overpower the liquidity effect, and the interest rate immediately starts to climb. Over time, as the income and price-level effects start to take hold, the interest rate rises even higher, and the eventual outcome is an interest rate that is substantially higher than the initial interest rate. The result shows clearly that increasing the money supply growth rate is not the answer to reducing interest rates; rather, money growth should be *reduced* to lower interest rates!

An important issue for economic policymakers is deciding which of these three scenarios is closest to reality. If a decline in interest rates is desired, then an increase in the money supply growth rate is called for when the liquidity effect dominates the other effects, as in panel (a). A decrease in money growth is appropriate when the other effects dominate the liquidity effect and expectations of inflation adjust rapidly, as in panel (c). If the other effects dominate the liquidity effect but expectations of inflation adjust only slowly, as in panel (b), then the decision to increase or decrease the money growth rate depends on whether policymakers care more about what happens in the short run or in the long run.

Mini-lecture

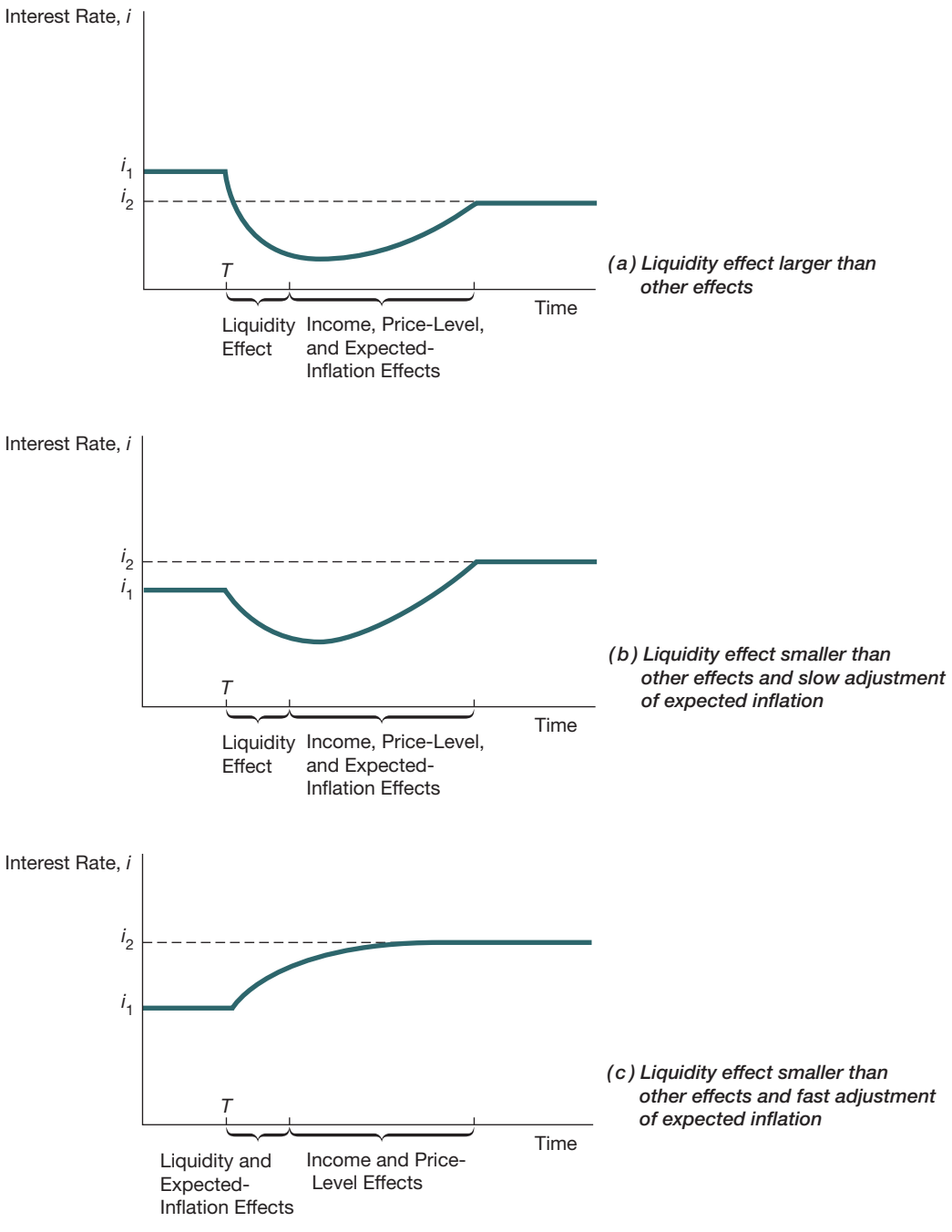


FIGURE 5-11 Response over Time to an Increase in Money Supply Growth

Each panel shows how interest rates respond over time to an increased rate of money supply growth, starting at time *T*.

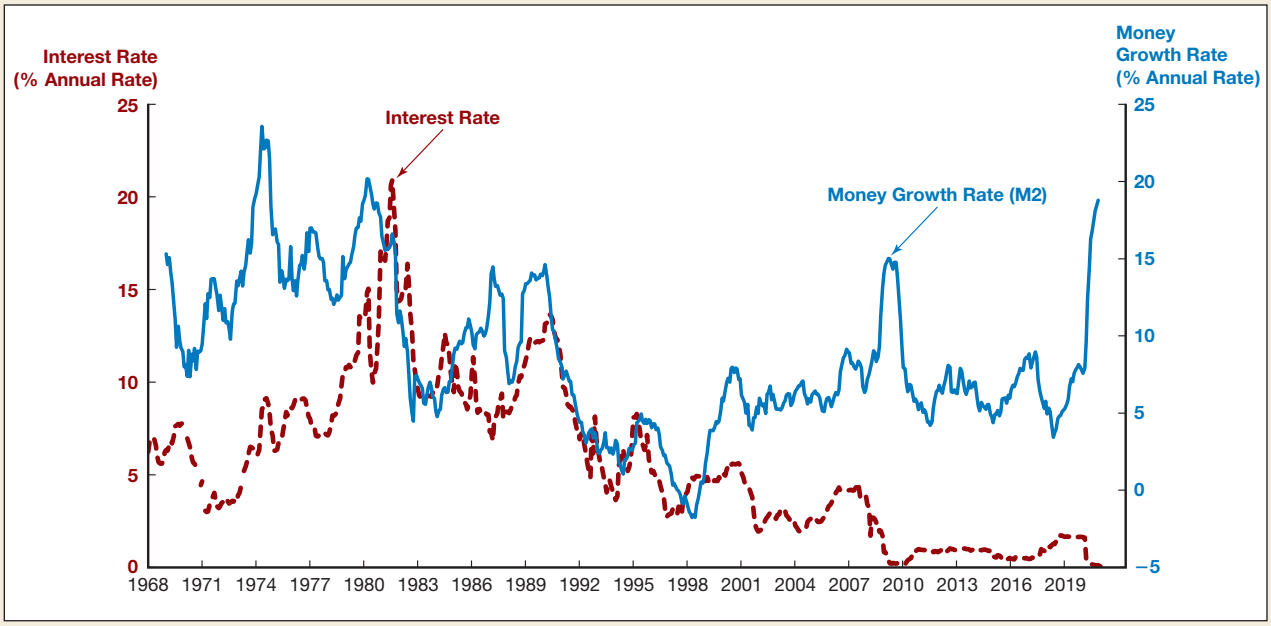


FIGURE 5-12 Money Growth (M2, Annual Rate) and Interest Rates (Three-Month Treasury Bills), 1968–2020

When the rate of money supply growth began to climb in the mid-1960s, interest rates rose, indicating that the liquidity effect was dominated by the price-level, income, and expected-inflation effects. By the early 1980s, both interest rates and money growth reached levels unprecedented in the post–World War II period.

Source: Statistics Canada CANSIM series V122531 and V41552796.

Which scenario is supported by the evidence? The relationship between interest rates and money growth from 1968 to 2020 is plotted in Figure 5-12. When the rate of money supply growth began to climb in the mid-1970s, interest rates rose, indicating that the liquidity effect was dominated by the price-level, income, and expected-inflation effects. By the early 1980s, interest rates reached levels unprecedented in the post–World War II period, as did the rate of money supply growth.

The scenario depicted in panel (a) of Figure 5-11 seems doubtful, and the case for lowering interest rates by raising the rate of money growth is much weakened. Looking back at Figure 5-5, which shows the relationship between interest rates and expected inflation, you should not find this too surprising. The rise in the rate of money supply growth in the 1960s and 1970s is matched by a large rise in expected inflation, while the fall in money growth from the early 1980s until the early 1990s is also associated with a decline in expected inflation. Both of these facts suggest that the expected inflation effect is dominant. It is the most plausible explanation for why interest rates rose in the face of higher money growth and then fell with the decline in money growth. However, Figure 5-12 does not really tell us which one of the two scenarios depicted in panels (b) and (c) of Figure 5-11 is more accurate. It depends critically on how fast people's expectations about inflation adjust. Research done using more sophisticated methods than just looking at a graph like Figure 5-12 indicates that increased money growth temporarily lowers short-term interest rates and thus conforms to the scenario in panel (b). ◆

SUMMARY

1. The theory of portfolio choice tells us that the quantity demanded of an asset is (a) positively related to wealth, (b) positively related to the expected return on the asset relative to alternative assets, (c) negatively related to the riskiness of the asset relative to alternative assets, and (d) positively related to the liquidity of the asset relative to alternative assets.
2. The supply and demand analysis for bonds provides one theory of how interest rates are determined. It predicts that interest rates will change when there is a change in demand caused by changes in income (or wealth), expected returns, risk, or liquidity, or when there is a change in supply caused by changes in the attractiveness of investment opportunities, the real cost of borrowing, or the government budget.
3. An alternative theory of how interest rates are determined is provided by the liquidity preference framework, which analyzes the supply of and demand for money. It shows that interest rates will change when the demand for money changes because of alterations in income or the price level, or when the supply of money changes.
4. An increase in the money supply has four possible effects on interest rates: the liquidity effect, the income effect, the price-level effect, and the expected-inflation effect. The liquidity effect indicates that a rise in money supply growth will lead to a decline in interest rates; the other effects work in the opposite direction. The evidence seems to indicate that the income, price-level, and expected-inflation effects dominate the liquidity effect such that an increase in the money supply growth rate leads to higher—rather than lower—interest rates.

MyLab Economics Make the grade with MyLab Economics: Select end-of-chapter questions can be found on MyLab Economics. You can practise them as often as you want, and most feature step-by-step guided instructions to help you find the right answer.

QUESTIONS

1. Explain why you would be more or less willing to buy a share of BlackBerry stock in the following situations:
   a. Your wealth falls.
   b. You expect the stock to appreciate in value.
   c. The bond market becomes more liquid.
   d. You expect gold to appreciate in value.
   e. Prices in the bond market become more volatile.
2. Explain why you would be more or less willing to buy a house under the following circumstances:
   a. You just inherited $100 000.
   b. Real estate commissions fall from 6% of the sales price to 5% of the sales price.
   c. You expect BlackBerry stock to double in value next year.
   d. Prices in the stock market become more volatile.
   e. You expect housing prices to fall.
3. Explain why you would be more or less willing to buy gold under the following circumstances:
   a. Gold again becomes acceptable as a medium of exchange.
   b. Prices in the gold market become more volatile.
   c. You expect inflation to rise, and gold prices tend to move with the aggregate price level.
   d. You expect interest rates to rise.

4. Explain why you would be more or less willing to buy long-term Air Canada bonds under the following circumstances:
   a. Trading in these bonds increases, making them easier to sell.
   b. You expect a bear market in stocks (stock prices are expected to decline).
   c. Brokerage commissions on stocks fall.
   d. You expect interest rates to rise.
   e. Brokerage commissions on bonds fall.
5. What will happen to the demand for Rembrandt paintings if the stock market undergoes a boom? Why?
6. Raphael observes that at the current level of interest rates there is an excess supply of bonds, and therefore he anticipates an increase in the price of bonds. Is Raphael correct?
7. Suppose Maria prefers to buy a bond with a 7% expected return and 2% standard deviation of its expected return, while Jennifer prefers to buy a bond with a 4% expected return and 1% standard deviation of its expected return. Can you tell if Maria is more or less risk-averse than Jennifer?
8. What will happen in the bond market if the government imposes a limit on the amount of daily transactions? Which characteristic of an asset would be affected?
9. How might a sudden increase in people's expectations of future real estate prices affect interest rates?
10. Suppose that many big corporations decide not to issue bonds, since it is now too costly to comply with new financial market regulations. Can you describe the expected effect on interest rates?
11. In the aftermath of the global economic crisis that started to take hold in 2008, Canadian government budget deficits increased dramatically, yet interest rates on Canada bonds fell sharply and stayed low for quite some time. Does this make sense? Why or why not?
12. Will there be an effect on interest rates if brokerage commissions on stocks fall? Explain your answer.
13. The governor of the Bank of Canada announces in a press conference that they will fight the higher inflation rate with a new anti-inflation program. Predict what will happen to interest rates if the public believes them.
14. Suppose that people in France decide to permanently increase their savings rate. Predict what will happen to the French bond market in the future. Can France expect higher or lower domestic interest rates?
15. Suppose you are in charge of the financial department of your company and you have to decide whether to borrow short or long term. Checking the news, you realize that the government is about to engage in a major infrastructure plan in the near future. Predict what will happen to interest rates. Will you advise borrowing short or long term?
16. Would fiscal policymakers ever have reason to worry about potentially inflationary conditions? Why or why not?
17. Why should a rise in the price level (but not in expected inflation) cause interest rates to rise when the nominal money supply is fixed?
18. If the next governor of the Bank of Canada has a reputation for advocating an even slower rate of money growth than the current governor, what will happen to interest rates? Discuss the possible resulting situations.
19. M1 money growth in the United States was about 15% in 2011 and 2012, and 10% in 2013. Over the same time period, the yield on three-month Treasury bills was close to 0%. Given these high rates of money growth, why did interest rates stay so low rather than increase? What does this say about the income, price-level, and expected-inflation effects?

APPLIED PROBLEMS

20. Suppose you visit with a financial adviser, and you are considering investing some of your wealth in one of three investment portfolios: stocks, bonds, or commodities. Your financial adviser provides you with the following table, which gives the probabilities of possible returns from each investment:
    a. Which investment should you choose to maximize your expected return: stocks, bonds, or commodities?
    b. If you are risk-averse and have to choose between the stock and the bond investments, which should you choose? Why?

Stocks		Bonds		Commodities	
Probability	**Return**	**Probability**	**Return**	**Probability**	**Return**
0.25	12%	0.6	10%	0.2	20%
0.25	10%	0.4	7.50%	0.25	12%
0.25	8%			0.25	6%
0.25	6%			0.25	4%
				0.05	0%

21. An important way in which the Bank of Canada decreases the money supply is by selling bonds to the public. Using a supply and demand analysis for bonds, show what effect this action has on interest rates. Is your answer consistent with what you would expect to find with the liquidity preference framework?

22. Using both the liquidity preference framework and the supply and demand for bonds framework, show why interest rates are procyclical (rising when the economy is expanding and falling during recessions).

23. Using both the supply and demand for bonds and liquidity preference frameworks, show how interest rates are affected when the riskiness of bonds rises. Are the results the same in the two frameworks?

24. The demand curve and supply curve for one-year discount bonds with a face value of $1000 are represented by the following equations:

B^d: Price $= -0.6 \times$ Quantity $+ 1140$

B^s: Price $=$ Quantity $+ 700$

a. What is the expected equilibrium price and quantity of bonds in this market?

b. Given your answer to part (a), what is the expected interest rate in this market?

25. The demand curve and supply curve for one-year discount bonds with a face value of $1000 are represented by the following equations:

B^d: Price $= -0.6 \times$ Quantity $+ 1140$

B^s: Price $=$ Quantity $+ 700$

Suppose that, as a result of monetary policy actions, the Bank of Canada sells 80 bonds that it holds. Assume that bond demand and money demand are held constant.

a. How does the Bank of Canada policy affect the bond supply equation?

b. Calculate the effect of the Bank of Canada's action on the equilibrium interest rate in this market.

DATA ANALYSIS PROBLEMS

1. Go to the CANSIM database and download monthly data from 1962:1 to 2021:1 for the three-month T-bill rate (series V122531).

a. What is the average nominal interest rate, i_t?

b. Calculate the change in i_t, $\Delta i_t = i_{t+1} - i_t$. What is the average change? The standard deviation?

2. **Real-Time Data Analysis** Go to the St. Louis Federal Reserve FRED database and find data on the M1 money supply (M1SL) and the 10-year U.S. Treasury bond rate (GS10). For the M1 money supply indicator, adjust the *units* setting to "Percent Change from Year Ago," and for both variables, adjust the *frequency* setting to "Quarterly." Download the data into a spreadsheet.

a. Create a scatter plot, with money growth on the horizontal axis and the 10-year Treasury rate on the vertical axis, from 2000:Q1 to the most recent quarter of data available. On the scatter plot, graph a fitted (regression) line of the data (there are several ways to do this; however, one particular chart layout has this option built in). Based on the fitted line, are the data consistent with the liquidity effect? Briefly explain.

b. Repeat part (a), but this time compare the contemporaneous money growth rate with the interest rate four quarters later. For example, create a scatter plot comparing money growth

from 2000:Q1 with the interest rate from 2001:Q1, and so on, up to the most recent pairwise data available. Compare your results to those obtained in part (a), and interpret the liquidity effect as it relates to the income, price-level, and expected-inflation effects.

c. Repeat part (a) again, except this time compare the contemporaneous money growth rate with the interest rate eight quarters later. For example, create a scatter plot comparing money growth from 2000:Q1 with the interest rate from 2002:Q1, and so on, up to the most recent pairwise data available. Assuming the liquidity and other effects are fully incorporated into the bond market after two years, what do your results imply about the overall effect of money growth on interest rates?

d. Based on your answers to parts (a) through (c), how do the actual data on money growth and interest rates compare to the three scenarios presented in Figure 5-11?

6 The Risk and Term Structure of Interest Rates

Learning Objectives

6.1 Identify and explain the three factors affecting the risk structure of interest rates.

6.2 List and explain the three theories of why interest rates vary across different maturities.

Preview

In our supply and demand analysis of interest-rate behaviour in Chapter 5, we examined the determination of just one interest rate. Yet we saw earlier that there are enormous numbers of bonds on which the interest rates can and do differ. In this chapter, we complete the interest-rate picture by examining the relationships of the various interest rates to one another. Understanding why interest rates differ from bond to bond can help businesses, banks, insurance companies, and private investors decide which bonds to purchase as investments and which ones to sell.

We first look at why bonds with the same term to maturity have different interest rates. The relationship among these interest rates is called the **risk structure of interest rates**, although risk, liquidity, and income tax rules all play a role in determining the risk structure. A bond's term to maturity also affects its interest rate, and the relationship among interest rates on bonds with different terms to maturity is called the **term structure of interest rates**. In this chapter, we examine the sources and causes of fluctuations in interest rates relative to one another and look at a number of theories that explain these fluctuations.

6.1 RISK STRUCTURE OF INTEREST RATES

LO 6.1 Identify and explain the three factors affecting the risk structure of interest rates.

Figure 6-1 shows the yields to maturity for several categories of Canadian long-term bonds from 1978 to 2020. It shows us two important features of interest-rate behaviour for bonds of the same maturity: Interest rates on different categories of bonds, although they generally move together, differ from one another in any given year, and the spread (or difference) between the interest rates varies over time. The interest rates on corporate bonds, for example, were higher than those on Canada bonds and provincial bonds. In addition, the spread between the interest rates on Baa corporate bonds (riskier than Aaa corporate bonds) and Canada bonds was very large during the 1980–1982 and 1990–1991 recessions, as well as during the recent Great Recession that followed the global financial crisis from 2007 to 2009. Which factors are responsible for these phenomena?

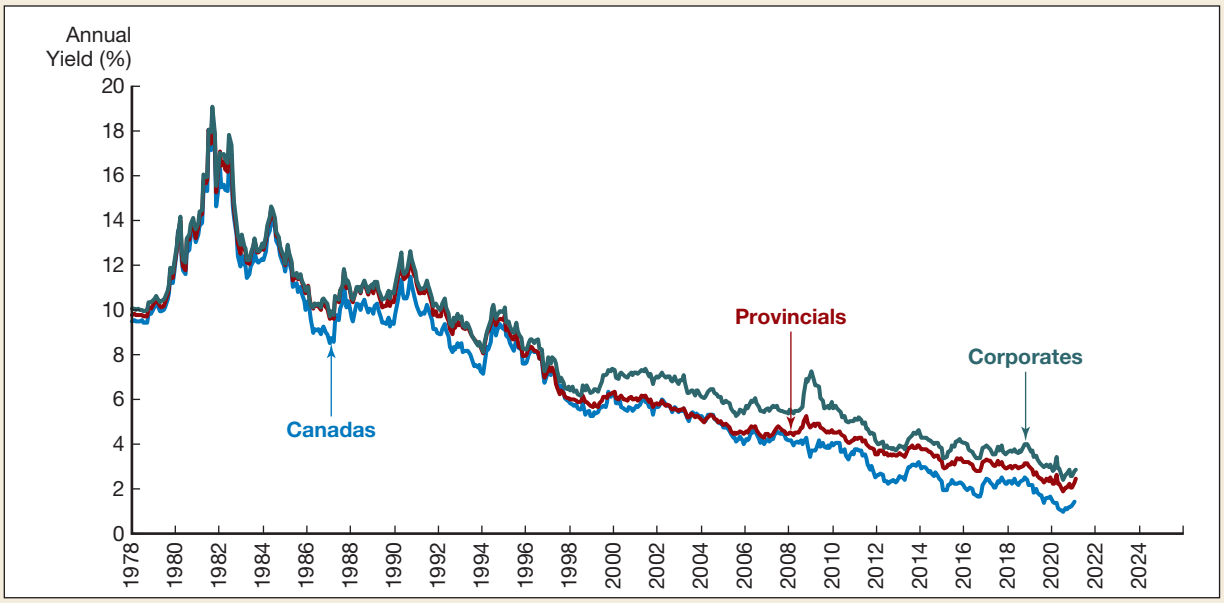


FIGURE 6-1 **Long-Term Bond Yields, 1978–2020**

Interest rates on different categories of bonds differ from one another in any given year, and the spread (or difference) between the interest rates varies over time.

Sources: Statistics Canada CANSIM series V122544, V122517 (extended with the average of provincial bond interest rates from Bloomberg), and V122518 (extended with the 30-year A-rated corporate bond interest rate from Bloomberg).

Default Risk

One attribute of a bond that influences its interest rate is its risk of **default**. Default occurs when the issuer of the bond is unable or unwilling to make interest payments when promised or pay off the face value when the bond matures. Corporations suffering big losses, such as airline and car rental companies during the coronavirus pandemic in 2020, might be more likely to suspend interest payments on their bonds. The default risk on their bonds would therefore be quite high. By contrast, Canadian government bonds have usually been considered to have no default risk because the federal government can always increase taxes or print money to pay off its obligations. Bonds like these with no default risk are called **default-free bonds**. The spread between interest rates on bonds with default risk and interest rates on default-free bonds, both of the same maturity, is called the **risk premium**. The risk premium indicates how much additional interest people must earn to be willing to hold the risky bond. Our supply and demand analysis of the bond market in Chapter 5 can be used to explain why a bond with default risk always has a positive risk premium and why a higher default risk means a larger risk premium.

To examine the effect of default risk on interest rates, let's look at the supply and demand diagrams for the default-free (Canadian government) and corporate long-term bond markets in Figure 6-2. To make the diagrams somewhat easier to read, let's assume that initially corporate bonds have the same default risk as Canada bonds. In this case, these two bonds have the same attributes (identical risk and maturity); their

Mini-lecture

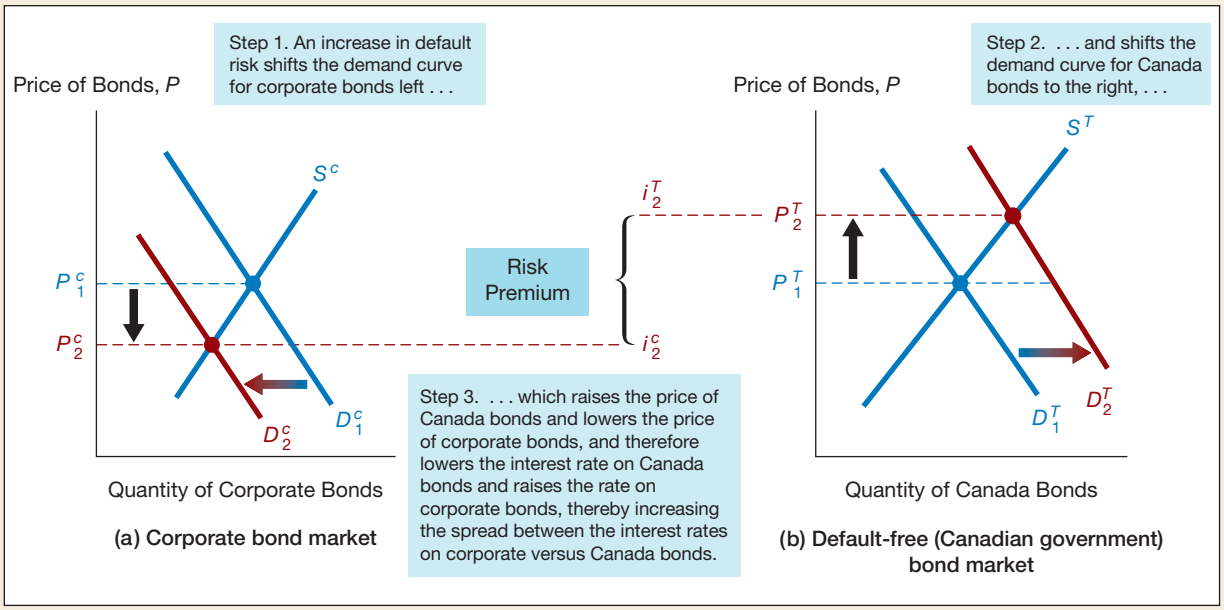

Step 1. An increase in default risk shifts the demand curve for corporate bonds left . . .
Price of Bonds, P
S^c
P^c_1
P^c_2
D^c_1
D^c_2
Quantity of Corporate Bonds
(a) Corporate bond market
Risk Premium
i^T_2
i^c_2
Step 3. . . . which raises the price of Canada bonds and lowers the price of corporate bonds, and therefore lowers the interest rate on Canada bonds and raises the rate on corporate bonds, thereby increasing the spread between the interest rates on corporate versus Canada bonds.
Step 2. . . . and shifts the demand curve for Canada bonds to the right, . . .
Price of Bonds, P
S^T
P^T_2
P^T_1
D^T_1
D^T_2
Quantity of Canada Bonds
(b) Default-free (Canadian government) bond market

FIGURE 6-2 Response to an Increase in Default Risk on Corporate Bonds

Initially, $P^c_1 = P^T_1$, $i^c_1 = i^T_1$, and the risk premium is zero. An increase in default risk on corporate bonds shifts the demand curve from D^c_1 to D^c_2; simultaneously, it shifts the demand curve for Canada bonds from D^T_1 to D^T_2. The equilibrium price for corporate bonds falls from P^c_1 to P^c_2, and the equilibrium interest rate on corporate bonds rises to i^c_2. In the Canadas market, the equilibrium bond price rises from P^T_1 to P^T_2, and the equilibrium interest rate falls to i^T_2. The brace indicates the difference between i^c_2 and i^T_2, the risk premium on corporate bonds. (Note that because P^c_2 is lower than P^T_2, i^c_2 is greater than i^T_2.)

equilibrium prices and interest rates will initially be equal ($P^c_1 = P^T_1$ and $i^c_1 = i^T_1$), and the risk premium on corporate bonds ($i^c_1 - i^T_1$) will be zero.

If the possibility of a default increases because a corporation begins to suffer large losses, the default risk on corporate bonds will increase and the expected return on these bonds will decrease. In addition, the corporate bond's return will be more uncertain. The theory of portfolio choice predicts that because the expected return on the corporate bond falls relative to the expected return on the default-free Canada bond while its relative riskiness rises, the corporate bond is less desirable (holding everything else equal), and demand for it will fall. Another way of thinking about this is that if you were an investor, you would want to hold (demand) a smaller amount of corporate bonds. The demand curve for corporate bonds in panel (a) of Figure 6-2 then shifts to the left, from D^c_1 to D^c_2.

At the same time, the expected return on default-free Canada bonds increases relative to the expected return on corporate bonds, while their relative riskiness declines. The Canada bonds thus become more desirable, and demand rises, as shown in panel (b) by the rightward shift in the demand curve for these bonds from D^T_1 to D^T_2.

As we can see in Figure 6-2, the equilibrium price for corporate bonds falls from P^c_1 to P^c_2, and since the bond price is negatively correlated to the interest rate, the equilibrium interest rate on corporate bonds rises to i^c_2. At the same time, however, the equilibrium price for the Canada bonds rises from P^T_1 to P^T_2 and the equilibrium interest rate falls to i^T_2.

The spread between the interest rates on corporate and default-free bonds—that is, the risk premium on corporate bonds—has risen from zero to $i_2^c - i_2^T$. We can now conclude that ***a bond with default risk will always have a positive risk premium, and an increase in its default risk will raise the risk premium.***

Because default risk is so important to the size of the risk premium, purchasers of bonds need to know whether a corporation is likely to default on its bonds. This information is provided by **credit-rating agencies**, investment advisory firms that rate the quality of corporate and municipal bonds in terms of their probability of default. Credit-rating agencies have become very controversial in recent years because of the role they played in the global financial crisis of 2007–2009 (see the FYI box "Conflicts of Interest at Credit-Rating Agencies and the Global Financial Crisis"). Table 6-1 provides the bond ratings and their descriptions as defined by the four largest credit-rating

TABLE 6-1 Bond Ratings by DBRS, Moody's, Standard & Poor's, and Fitch

	Rating			
DBRS	**Moody's**	**S&P**	**Fitch**	**Definitions**
AAA	Aaa	AAA	AAA	Prime Maximum Safety
AA high	Aa1	AA1	AA1	High Grade, High Quality
AA	Aa2	AA	AA	
AA low	Aa3	AA2	AA2	
A high	A1	A1	A1	Upper Medium Grade
A	A2	A	A	
A low	A3	A2	A2	
BBB high	Baa1	BBB1	BBB1	Lower Medium Grade
BBB	Baa2	BBB	BBB	
BBB low	Baa3	BBB2	BBB2	
BB high	Ba1	BB1	BB1	Noninvestment Grade
BB	Ba2	BB	BB	Speculative
BB low	Ba3	BB2	BB2	
B high	B1	B1	B1	Highly Speculative
B	B2	B	B	
B low	B3	B2	B2	
CCC	Caa1	CCC1	CCC	Substantial Risk
CC	Caa2	CCC	—	In Poor Standing
C	Caa3	CCC2	—	
—	Ca	—	—	Extremely Speculative
—	C	—	—	May Be in Default
—	—	—	DDD	Default
—	—	—	DD	—
D	—	D	D	

FYI **Conflicts of Interest at Credit-Rating Agencies and the Global Financial Crisis**

Debt ratings play a major role in the pricing of debt securities and in the regulatory process. Conflicts of interest arose for credit-rating agencies in the years leading up to the global financial crisis when these agencies advised clients on how to structure complex financial instruments that paid out cash flows from subprime mortgages. At the same time, the agencies were rating these identical products, leading to the potential for severe conflicts of interest. Specifically, the large fees they earned from advising clients on how to structure products that the agencies themselves were rating meant they did not have sufficient incentives to make sure their ratings were accurate.

When housing prices began to fall and subprime mortgages began to default, it became crystal clear that the ratings agencies had done a terrible job of assessing the risk in the subprime products they had helped to structure. Many AAA-rated products had to be downgraded over and over again until they reached junk status. The resulting massive losses on these assets were one reason why so many financial institutions that were holding them got into trouble. Indeed, some commentators have cited credit-rating agencies as among the chief villains that should be blamed for the global financial crisis.

agencies in Canada: Dominion Bond Rating Service (DBRS), Moody's Investor Service, Standard and Poor's Corporation, and Fitch Ratings. Bonds with relatively low risk of default are called *investment-grade* securities and have a rating of Baa (or BBB) and above. Bonds with ratings below Baa (or BBB) have higher default risk and have been aptly dubbed speculative-grade or **junk bonds**. Because these bonds always have higher interest rates than investment-grade securities, they are also referred to as high-yield bonds.

Next let's look at Figure 6-1 at the beginning of the chapter and see if we can explain the relationship between interest rates on corporate and those on Canada bonds. Corporate bonds always have higher interest rates than Canada bonds because they always have some risk of default, whereas Canada bonds do not. Because corporate bonds have a greater default risk than Canada bonds, their risk premium is greater, and the corporate bond rate therefore always exceeds the Canada bond rate. We can use the same analysis to explain the huge jump in the risk premium on corporate bond rates during the 1980–1982, 1990–1991, and 2000 recessions. The recession periods saw a very high rate of business failures and defaults. As we would expect, these factors led to a substantial increase in the default risk for bonds issued by vulnerable corporations, and the risk premium for corporate bonds reached unprecedentedly high levels.

APPLICATION The Coronavirus Pandemic and the Baa–Treasury Spread in the United States

When the coronavirus officially became a pandemic in March 2020, the resulting lockdowns and social distancing dealt a knockout blow to the world economy. Many investors began to doubt the financial health of corporations with low credit

ratings such as Baa. The perceived increase in default risk for Baa bonds made them less desirable at any given price, decreased the quantity demanded, and shifted the demand curve for Baa bonds to the left. As shown in panel (a) of Figure 6-2, the interest rate on Baa bonds should have risen, which is indeed what happened. Interest rates on Baa bonds jumped by 70 basis points (0.70 percentage points), from 3.6% in February 2020 to 4.3% in March 2020. At the same time, the increase in perceived default risk for Baa bonds made default-free U.S. Treasury bonds relatively more attractive and shifted the demand curve for these securities to the right—an outcome described by some analysts as a "flight to quality." Just as our analysis in Figure 6-2 predicts, interest rates on Treasury bonds fell by 60 basis points, from 1.50% in February 2020 to 0.9% in March 2020. In just one month, the spread between interest rates on Baa and Treasury bonds rose by 140 basis points, from 2.1% before the pandemic to 3.40% in March 2020, when the pandemic spread worldwide. ◆

Liquidity

Another attribute of a bond that influences its interest rate is its liquidity. As we learned in Chapter 4, a liquid asset is one that can be quickly and cheaply converted into cash if the need arises. The more liquid an asset is, the more desirable it is (holding everything else constant). Canada bonds are the most liquid of all long-term bonds; because they are so widely traded, they are the easiest to sell quickly, and the cost of selling them is low. Corporate bonds are not as liquid because fewer bonds for any one corporation are traded; thus it can be costly to sell these bonds in an emergency because it might be hard to find buyers quickly.

How does the reduced liquidity of the corporate bonds affect their interest rates relative to the interest rate on Canada bonds? We can use supply and demand analysis with the same figure that was used to analyze the effect of default risk, Figure 6-2, to show that the lower liquidity of corporate bonds relative to Canada bonds increases the spread between the interest rates on these two bonds.

Let's start the analysis by assuming that, initially, corporate and Canada bonds are equally liquid and all their other attributes are the same. As shown in Figure 6-2, their equilibrium prices and interest rates will initially be equal: $P_1^c = P_1^T$ and $i_1^c = i_1^T$. If the corporate bond becomes less liquid than the Canada bond because it is less widely traded, then (as the theory of portfolio choice indicates) demand for it will fall, shifting its demand curve leftward from D_1^c to D_2^c, as in panel (a). The Canada bond now becomes relatively more liquid in comparison with the corporate bond, so its demand curve shifts rightward from D_1^T to D_2^T, as in panel (b). The shifts in the curves in Figure 6-2 show that the price of the less liquid corporate bond falls and its interest rate rises, while the price of the more liquid Canada bond rises and its interest rate falls.

The result is that the spread between the interest rates on the two bond types rises. Therefore, the differences between interest rates on corporate bonds and Canada bonds (that is, the risk premiums) reflect not only the corporate bonds' default risk but also their lesser liquidity. This is why a risk premium should more accurately be called a "risk and liquidity premium," but convention dictates the label *risk premium*.

Income Tax Considerations

In Canada, coupon payments on fixed-income securities are taxed as ordinary income in the year they are received. In some other countries, however, certain government bonds are not taxable. In the United States, for example, interest payments on municipal bonds are exempt from federal income taxes, and these bonds have had lower interest rates than U.S. Treasury bonds for most of the past 100 years. How does taxation affect the interest rate on bonds?

Let's imagine that your income is high enough to put you in a 40% income tax bracket, where for every extra dollar of income you earn you have to pay 40 cents to the government. If you own a $1000-face-value taxable bond that sells for $1000 and has a coupon payment of $100, you get to keep only $60 of the payment after taxes. Although the bond has a 10% interest rate, you actually earn only 6% after taxes.

Suppose, however, that you put your savings into a $1000-face-value tax-exempt bond that sells for $1000 and pays only $80 in coupon payments. Its interest rate is only 8%, but because it is a tax-exempt security, you pay no taxes on the $80 coupon payment, so you earn 8% after taxes. Clearly, you earn more on the tax-exempt bond after taxes, so you are willing to hold the bond even though it has a lower interest rate than the taxable bond.

Another way of understanding why tax-exempt bonds have lower interest rates than taxable bonds is to use the supply and demand analysis depicted in Figure 6-3. We initially assume that municipal and Treasury bonds in the United States have identical attributes and so have the same bond prices as drawn in the figure, $P_1^m = P_1^T$, and the same interest rates. Once the municipal bonds are given a tax advantage that raises their after-tax expected return relative to U.S. Treasury bonds and makes them more desirable, demand for them rises, and their demand curve shifts to the right, from D_1^m to D_2^m. Their equilibrium bond price then rises from P_1^m to P_2^m, and their equilibrium interest rate falls. By contrast, U.S. Treasury bonds have now become less desirable relative to U.S. municipal bonds; demand for Treasury bonds decreases, and D_1^T shifts to the left to D_2^T. The U.S. Treasury bond price falls from P_1^T to P_2^T, and the interest rate rises. The federal income tax exemption, which leads to a higher expected return for U.S. municipal bonds relative to U.S. Treasury bonds, explains why U.S. municipal bonds can have interest rates below those of U.S. Treasury bonds.[1]

Summary

The risk structure of interest rates (the relationships among interest rates on bonds with the same maturity) is explained by three factors: default risk, liquidity, and the income tax treatment of a bond's interest payments. As a bond's default risk increases, the risk premium on that bond (the spread between its interest rate and the interest rate on a default-free Canada bond) rises. The greater liquidity of Canada bonds also explains why their interest rates are lower than those on less liquid bonds. If a bond has a favourable tax treatment, as do municipal bonds in the United States, its interest rate will be lower.

[1]In contrast to corporate bonds, U.S. Treasury bonds are exempt from state and local income taxes. Using the analysis in the text, you should be able to show that this feature of U.S. Treasury bonds provides an additional reason why interest rates on corporate bonds are higher than those on U.S. Treasury bonds.

Mini-lecture

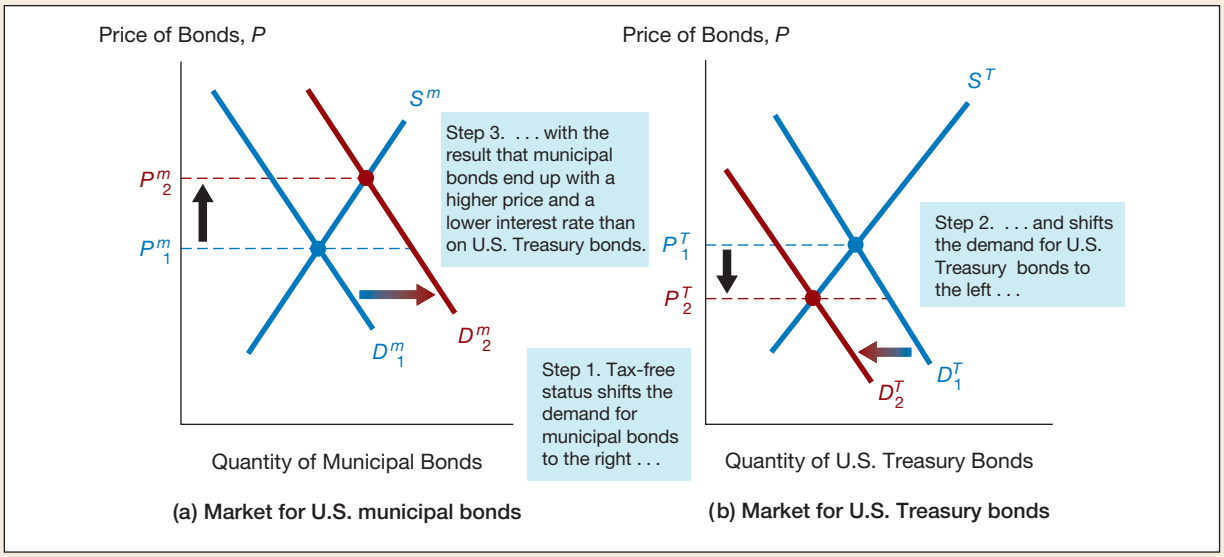

Price of Bonds, P
S^m
P^m_2
P^m_1
D^m_1
D^m_2
Step 3. . . . with the result that municipal bonds end up with a higher price and a lower interest rate than on U.S. Treasury bonds.
Step 1. Tax-free status shifts the demand for municipal bonds to the right . . .
Quantity of Municipal Bonds
(a) Market for U.S. municipal bonds
Price of Bonds, P
S^T
P^T_1
P^T_2
D^T_1
D^T_2
Step 2. . . . and shifts the demand for U.S. Treasury bonds to the left . . .
Quantity of U.S. Treasury Bonds
(b) Market for U.S. Treasury bonds

FIGURE 6-3 Interest Rates on Municipal and Treasury Bonds in the United States

When a municipal bond is given tax-free status, demand for the municipal bond shifts rightward from D^m_1 to D^m_2 and demand for the U.S. Treasury bond shifts leftward from D^T_1 to D^T_2. The equilibrium price of the municipal bond rises from P^m_1 to P^m_2, so its interest rate falls, while the equilibrium price of the U.S. Treasury bond falls from P^T_1 to P^T_2 and its interest rate rises. The result is that municipal bonds end up with lower interest rates than U.S. Treasury bonds.

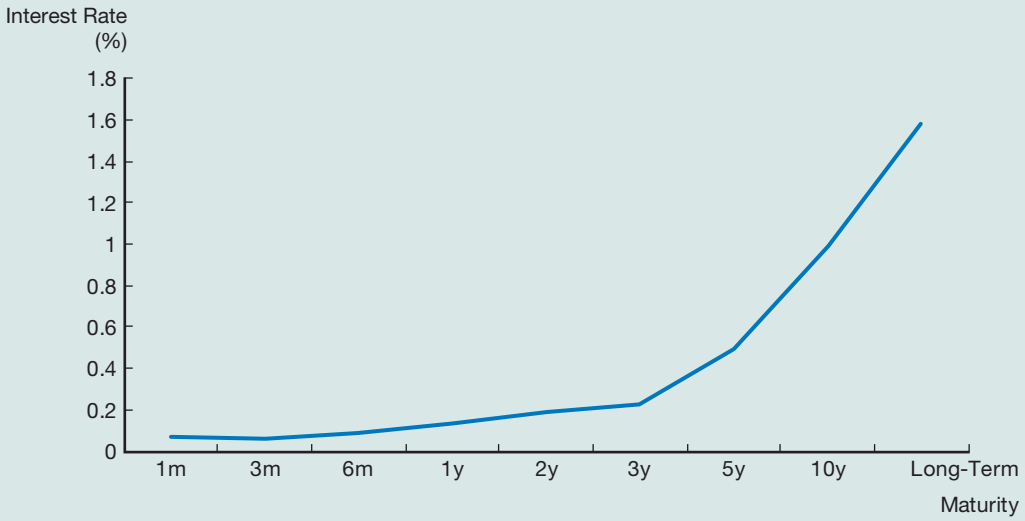
6.2 TERM STRUCTURE OF INTEREST RATES

LO 6.2 List and explain the three theories of why interest rates vary across different maturities.

We have seen how risk, liquidity, and tax considerations (collectively embedded in the risk structure) can influence interest rates. Another factor that influences the interest rate on a bond is its term to maturity: Bonds with identical risk, liquidity, and tax characteristics may have different interest rates because their times remaining to maturity are different. A plot of the yields on bonds with differing terms to maturity but the same risk, liquidity, and tax considerations is called a **yield curve**, and it describes the term structure of interest rates for particular types of bonds, such as government bonds. The Following the Financial News box shows a yield curve for Canada securities. Yield curves can be classified as upward-sloping, flat, or downward-sloping (the last sort is often referred to as an **inverted yield curve**). When yield curves slope upward, which is the most common case, long-term interest rates are above short-term interest rates; when yield curves are flat, short- and long-term interest rates are the same; and when yield curves are inverted, long-term interest rates are below short-term interest rates. Yield curves can also have more complicated shapes in which they first slope up and then down, or vice versa. Why does the yield curve usually slope upward but sometimes take on other shapes?

Following the Financial News Yield Curves

Many newspapers and Internet sites such as www.finance.yahoo.com publish a daily plot of the yield curves for government securities. An example for February 3, 2021, is presented here. The numbers on the vertical axis indicate the interest rate for the Canada security, with the maturity term given on the horizontal axis, with "m" denoting "month" and "y" denoting "year."



In addition to explaining why yield curves take on different shapes at different times, a good theory of the term structure of interest rates must explain the following three important empirical facts:

1. As we see in Figure 6-4, interest rates on bonds of different maturities move together over time.
2. When short-term interest rates are low, yield curves are more likely to have an upward slope; when short-term interest rates are high, yield curves are more likely to slope downward and be inverted.
3. Yield curves almost always slope upward, as appears in the Following the Financial News box.

Three theories have been put forward to explain the term structure of interest rates—that is, the relationships among interest rates on bonds of different maturities reflected in yield curve patterns—(1) the expectations theory, (2) the segmented markets theory, and (3) the liquidity premium theory. Each of these theories is described in the following sections. The expectations theory does a good job of explaining the first two facts on our list but not the third. The segmented markets theory can account for fact 3 but not the other two facts, which are well explained by the expectations theory. Because each theory explains facts that the other cannot, a natural way of seeking a better understanding of the term structure is to combine the features of both theories, which leads us to the liquidity premium theory, a theory that explains all three facts.

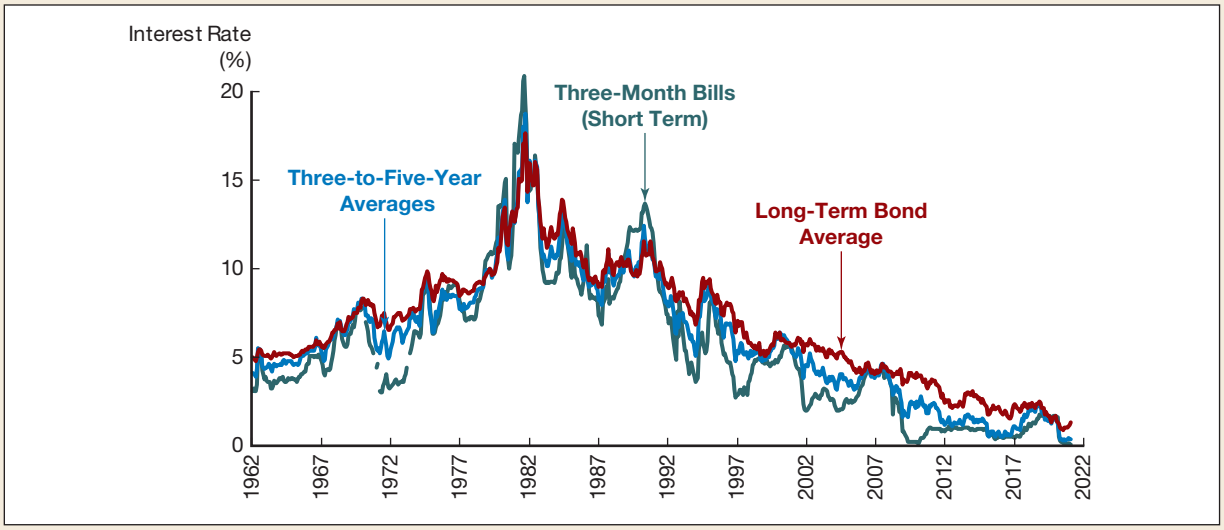


FIGURE 6-4 Movements over Time of Interest Rates on Government of Canada Bonds with Different Maturities

Interest rates on bonds of different maturities move together over time.

Source: Statistics Canada CANSIM series V122531, V122485, and V122487.

If the liquidity premium theory does a better job of explaining the facts and is hence the most widely accepted theory, why do we spend time discussing the other two theories? There are two reasons. First, the ideas in these two theories lay the groundwork for the liquidity premium theory. Second, it is important to see how economists modify theories to improve them when they find that the predicted results are inconsistent with the empirical evidence.

Expectations Theory

The **expectations theory** of the term structure states the following commonsense proposition: The interest rate on a long-term bond will equal the average of the short-term interest rates that people expect to occur over the life of the long-term bond. For example, if people expect that short-term interest rates will be 10%, on average, over the coming five years, the expectations theory predicts that the interest rate on bonds with five years to maturity will be 10% too. If short-term interest rates are expected to rise even higher after this five-year period, so that the average short-term interest rate over the coming 20 years is 11%, then the interest rate on 20-year bonds will equal 11% and will be higher than the interest rate on five-year bonds. Thus the expectations theory predicts that interest rates on bonds of different maturities differ because short-term interest rates are expected to have different values at future dates.

The key assumption behind this theory is that buyers of bonds do not prefer bonds of one maturity over another, so they will not hold any quantity of a bond if its expected

return is less than that of another bond with a different maturity. Bonds that have this characteristic are said to be *perfect substitutes*. In practice, this means that if bonds with different maturities are perfect substitutes, then the expected returns on those bonds must be equal.

To see how the assumption that bonds with different maturities are perfect substitutes leads to the expectations theory, let us consider the following two investment strategies:

1. Purchase a one-year bond, and when it matures in one year, purchase another one-year bond.
2. Purchase a two-year bond and hold it until maturity.

Because both strategies must have the same expected return, the interest rate on the two-year bond must equal the average of the two one-year interest rates. For example, let's say that the current interest rate on the one-year bond is 9%, and you expect the interest rate on the one-year bond next year to be 11%. If you pursue the first strategy of buying the two one-year bonds, the expected return over the two years will average out to be $(9\% + 11\%)/2 = 10\%$ per year. You will be willing to hold both the one- and two-year bonds only if the expected return per year on the two-year bond equals this return. Therefore, the interest rate on the two-year bond must equal 10%, the average interest rate on the two one-year bonds.

We can make this argument more general. For an investment of \$1, consider the choice of holding, for two periods, a two-period bond or two one-period bonds. Using the definitions

$$\begin{aligned} i_t &= \text{today's (time } t\text{) interest rate on a one-period bond} \\ i^e_{t+1} &= \text{interest rate on a one-period bond expected for next period (time } t+1\text{)} \\ i_{2t} &= \text{today's (time } t\text{) interest rate on the two-period bond} \end{aligned}$$

the expected return from investing \$1 in the two-period bond and holding it for the two periods can be calculated as

$$(1 + i_{2t})(1 + i^e_{2t}) - 1 = 1 + 2i_{2t} + (i_{2t})^2 - 1 = 2i_{2t} + (i_{2t})^2$$

After the second period, the \$1 investment is worth $(1 + i_{2t})(1 + i_{2t})$. Subtracting the \$1 initial investment from this amount and dividing by the initial \$1 investment gives the rate of return calculated in the previous equation. Because $(i_{2t})^2$ is extremely small—if $i_{2t} = 10\% = 0.10$, then $(i_{2t})^2 = 0.01$—we can simplify the expected return for holding the two-period bond for the two periods to

$$2i_{2t}$$

With the other strategy, in which one-period bonds are bought, the expected return on the \$1 investment over the two periods is

$$(1 + i_t)(1 + i^e_{t+1}) - 1 = 1 + i_t + i^e_{t+1} + i_t(i^e_{t+1}) - 1 = i_t + i^e_{t+1} + i_t(i^e_{t+1})$$

This calculation is derived by recognizing that, after the first period, the \$1 investment becomes $1 + i_t$, and this is reinvested in the one-period bond for the next period, yielding an amount $(1 + i_t)(1 + i^e_{t+1})$. Then, subtracting the \$1 initial investment from this amount and dividing by the initial investment of \$1 gives the expected return for

the strategy of holding one-period bonds for the two periods. Because $i_t(i^e_{t+1})$ is also extremely small—if $i_t = i^e_{t+1} = 0.10$, then $i_t(i^e_{t+1}) = 0.01$—we can simplify this to

$$i_t + i^e_{t+1}$$

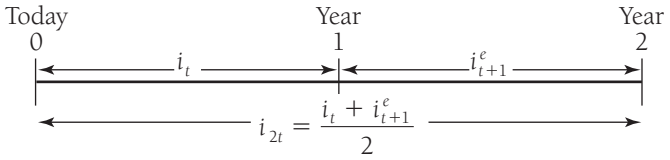
Both bonds will be held only if these expected returns are equal—that is, when

$$2i_{2t} = i_t + i^e_{t+1}$$

Solving for i_{2t} in terms of the one-period rates, we have

$$i_{2t} = \frac{i_t + i^e_{t+1}}{2} \tag{1}$$

which tells us that the two-period rate must equal the average of the two one-period rates. Graphically, this can be shown as



We can conduct the same steps for bonds with longer maturities so that we can examine the whole term structure of interest rates. By doing so, we find that the interest rate i_{nt} on an n-period bond must be

$$i_{nt} = \frac{i_t + i^e_{t+1} + i^e_{t+2} + \ldots + i^e_{t+(n-1)}}{n} \tag{2}$$

Equation 2 states that the n-period interest rate equals the average of the one-period interest rates expected to occur over the n-period life of the bond. This is a restatement of the expectations theory in more precise terms.[2]

A simple numerical example might help clarify the expectations theory presented in Equation 2. If the one-year interest rates over the next five years are expected to be 5%, 6%, 7%, 8%, and 9%, Equation 2 indicates that the interest rate on the two-year bond will be

$$\frac{5\% + 6\%}{2} = 5.5\%$$

On the five-year bond, it will be

$$\frac{5\% + 6\% + 7\% + 8\% + 9\%}{5} = 7\%$$

By doing similar calculations for the one-, three-, and four-year interest rates, you should be able to verify that the one- to five-year interest rates are 5.0%, 5.5%, 6.0%, 6.5%, and 7.0%, respectively. Thus we see that the rising trend in expected short-term interest rates produces an upward-sloping yield curve along which interest rates rise as maturity lengthens.

[2]The analysis here has been conducted for discount bonds. Formulas for interest rates on coupon bonds would differ slightly from those used here but would convey the same principle.

The expectations theory is an elegant theory that explains why the term structure of interest rates (as represented by yield curves) changes at different times. When the yield curve is upward-sloping, the expectations theory suggests that short-term interest rates are expected to rise in the future, as we have seen in our numerical example. In this situation, in which the long-term rate is currently higher than the short-term rate, the average of future short-term rates is expected to be higher than the current short-term rate, which can occur only if short-term interest rates are expected to rise. This result is depicted in our numerical example. When the yield curve is inverted (slopes downward), the average of future short-term interest rates is expected to be lower than the current short-term rate, implying that short-term interest rates are expected to fall, on average, in the future. Only when the yield curve is flat does the expectations theory suggest that short-term interest rates are not expected to change, on average, in the future.

The expectations theory also explains fact 1, which states that interest rates on bonds with different maturities move together over time. Historically, short-term interest rates have shown the following trend: If they increase today, they will tend to be higher in the future. Hence a rise in short-term rates will raise people's expectations of future higher short-term rates. Because long-term rates are the average of expected future short-term rates, a rise in short-term rates will also raise long-term rates, causing short- and long-term rates to move together.

The expectations theory also explains fact 2, which states that yield curves tend to have an upward slope when short-term interest rates are low and be inverted when short-term rates are high. When short-term rates are low, people generally expect them to rise to some normal level in the future, and the average of future expected short-term rates is high relative to the current short-term rate. Therefore, long-term interest rates will be substantially higher than current short-term rates, and the yield curve will have an upward slope. Conversely, if short-term rates are high, people usually expect them to come back down. Long-term rates will then drop below short-term rates because the average of expected future short-term rates will be lower than current short-term rates, and the yield curve will slope downward and become inverted.[3]

The expectations theory is an attractive theory because it provides a simple explanation of the behaviour of the term structure, but unfortunately it has a major shortcoming: It cannot explain fact 3, which states that yield curves usually slope upward. The typical upward slope of yield curves implies that short-term interest rates are usually expected to rise in the future. In practice, short-term interest rates are just as likely to fall as they are to rise, and so the expectations theory suggests that the typical yield curve should be flat rather than upward-sloping.

Segmented Markets Theory

As the name suggests, the **segmented markets theory** of the term structure sees markets for different-maturity bonds as completely separate and segmented. The interest rate on a bond of a particular maturity is then determined by the supply of and

[3]The expectations theory explains another important fact about the relationship between short-term and long-term interest rates. As you can see in Figure 6-4, short-term interest rates are more volatile than long-term rates. If interest rates are *mean-reverting*—that is, if they tend to head back down after they reach unusually high levels or go back up after they fall to unusually low levels—then an average of these short-term rates must necessarily have less volatility than the short-term rates themselves. Because the expectations theory suggests that the long-term rate will be equal to an average of future short-term rates, it implies that the long-term rate will have less volatility than short-term rates.

demand for that bond and is not affected by expected returns on other bonds with other maturities.

The key assumption of the segmented markets theory is that bonds of different maturities are not substitutes at all, and so the expected return from holding a bond of one maturity has no effect on the demand for a bond of another maturity. This theory of the term structure is the opposite extreme of the expectations theory, which assumes that bonds of different maturities are perfect substitutes.

The argument for why bonds of different maturities are not substitutes is that investors have very strong preferences for bonds of one maturity as opposed to another, so they are concerned only with the expected returns on bonds of the maturity they prefer. This might be because they have a particular holding period in mind, and if they match the maturity of the bond to the desired holding period, they can obtain a certain return with no risk at all.[4] (We saw in Chapter 4 that if the term to maturity equals the holding period, then the return is known for certain because it equals the yield exactly, and no interest-rate risk exists.) For example, people who are limited by a short holding period will prefer to hold short-term bonds. Conversely, if you are putting away funds for your young child to go to college in the future, your desired holding period will be much longer, and you will want to hold longer-term bonds.

In the segmented markets theory, differing yield curve patterns are accounted for by supply and demand differences associated with bonds of different maturities. If, as seems sensible, risk-averse investors have short desired holding periods and generally prefer bonds with shorter maturities that have less interest-rate risk, the segmented markets theory can explain fact 3, which states that yield curves typically slope upward. Because the demand for long-term bonds is typically relatively lower than that for short-term bonds, long-term bonds will have lower prices and higher interest rates, and hence the yield curve will typically slope upward.

Although the segmented markets theory can explain why yield curves usually tend to slope upward, it has a major flaw in that it cannot explain facts 1 and 2. First, because it views the market for bonds of different maturities as completely segmented, there is no reason that a rise in the interest rate on a bond of one maturity would affect the interest rate on a bond of another maturity. Therefore, the segmented markets theory cannot explain why interest rates on bonds of different maturities tend to move together (fact 1). Second, because it is not clear how the demand for and supply of short- versus long-term bonds change with the level of short-term interest rates, the theory does not explain why yield curves tend to slope upward when short-term interest rates are low and to be inverted when short-term interest rates are high (fact 2).

Because each of our two theories explains empirical facts that the other cannot, a logical step is to combine the theories, which leads us to the liquidity premium theory.

Liquidity Premium and Preferred Habitat Theories

The **liquidity premium theory** of the term structure states that the interest rate on a long-term bond will equal an average of short-term interest rates expected to occur over the life of the long-term bond plus a liquidity premium (also referred to as a term premium) that responds to supply and demand conditions for that bond.

[4]The statement that there is no uncertainty about the return if the term to maturity equals the holding period is literally true only for a discount bond. For a coupon bond with a long holding period, some risk exists because coupon payments must be reinvested before the bond matures. Our analysis here is thus being conducted for discount bonds. However, the gist of the analysis remains the same for coupon bonds because the amount of risk from reinvestment is small when coupon bonds have the same term to maturity as the holding period.

The liquidity premium theory's key assumption is that bonds of different maturities are substitutes, which means that the expected return on one bond *does* influence the expected return on a bond of a different maturity. However, the theory allows investors to prefer one bond maturity over another. In other words, bonds of different maturities are assumed to be substitutes but not *perfect* substitutes. Investors tend to prefer shorter-term bonds because these bonds bear less interest-rate risk. For this reason, investors must be offered a positive liquidity premium to induce them to hold longer-term bonds. Thus the expectations theory is modified by adding a positive liquidity premium to the equation that describes the relationship between long- and short-term interest rates. The liquidity premium theory is written as

$$i_{nt} = \frac{i_t + i^e_{t+1} + i^e_{t+2} + \ \ldots \ + i^e_{t+(n-1)}}{n} + l_{nt} \tag{3}$$

where l_{nt} is the liquidity (term) premium for the n-period bond at time t, which is always positive and rises with the term to maturity of the bond, n.

Closely related to the liquidity premium theory is the **preferred habitat theory**, which takes a somewhat less direct approach to modifying the expectations hypothesis but comes to a similar conclusion. It assumes that investors have a preference for bonds of one maturity over bonds of another—a particular bond maturity ("preferred habitat") in which they prefer to invest. Because they prefer bonds of one maturity over bonds of another, they are willing to buy bonds that do not have the preferred maturity (habitat) only if those bonds earn a somewhat higher expected return. Because risk-averse investors are likely to prefer the habitat of short-term bonds over that of longer-term bonds, they are willing to hold long-term bonds only if they have higher expected returns. This reasoning leads to the same Equation 3 implied by the liquidity premium theory, with a term premium that typically rises with maturity.

The relationship between the expectations theory and the liquidity premium and preferred habitat theories is shown in Figure 6-5. There we see that because the liquidity premium is always positive and typically grows as the term to maturity increases, the yield curve implied by the liquidity premium theory is always above the yield curve implied by the expectations theory and generally has a steeper slope. (For simplicity, we are assuming that the expectations theory yield curve is flat.)

A simple numerical example, similar to the one we used for the expectations hypothesis, further clarifies the liquidity premium and preferred habitat theories given in Equation 3. Again suppose that the one-year interest rates over the next five years are expected to be 5%, 6%, 7%, 8%, and 9%, while investors' preferences for holding short-term bonds means that the liquidity premiums for one- to five-year bonds are 0%, 0.25%, 0.5%, 0.75%, and 1.0%, respectively. Equation 3 then indicates that the interest rate on the two-year bond would be

$$\frac{5\% + 6\%}{2} + 0.25\% = 5.75\%$$

For the five-year bond, it would be

$$\frac{5\% + 6\% + 7\% + 8\% + 9\%}{5} + 1\% = 8\%$$

By doing similar calculations for the one-, three-, and four-year interest rates, you should be able to verify that the one- to five-year interest rates are 5.0%, 5.75%, 6.5%, 7.25%, and

Mini-lecture

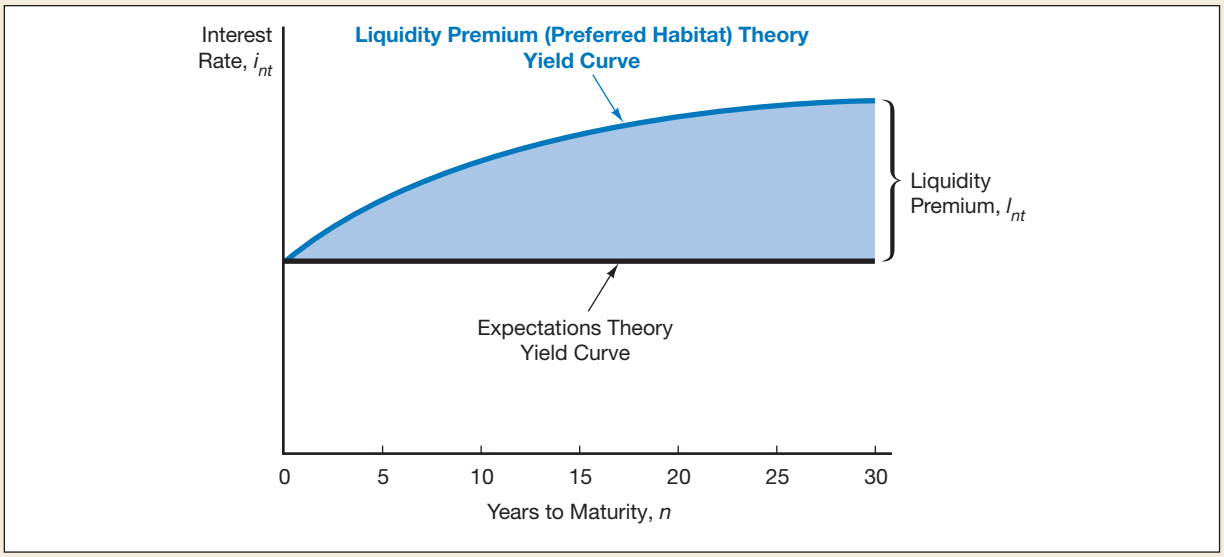


FIGURE 6-5 The Relationship Between the Liquidity Premium (Preferred Habitat) and Expectations Theory

Because the liquidity premium is always positive and grows as the term to maturity increases, the yield curve implied by the liquidity premium and preferred habitat theories is always above the yield curve implied by the expectations theory and has a steeper slope. For simplicity, the yield curve implied by the expectations theory shown here assumes unchanging future one-year interest rates.

8.0%, respectively. Comparing these findings with those for the expectations theory, we see that the liquidity premium and preferred habitat theories produce yield curves that slope more steeply upward because of investors' preferences for short-term bonds.

Let's see if the liquidity premium and preferred habitat theories are consistent with all three empirical facts we have discussed. They explain fact 1, which states that interest rates on different-maturity bonds move together over time: A rise in short-term interest rates indicates that short-term interest rates will, on average, be higher in the future, and the first term in Equation 3 then implies that long-term interest rates will rise along with them.

The liquidity and preferred habitat theories also explain why yield curves tend to have an especially steep upward slope when short-term interest rates are low and to be inverted when short-term rates are high (fact 2). Because investors generally expect short-term interest rates to rise to some normal level when they are low, the average of future expected short-term rates will be high relative to the current short-term rate. With the additional boost of a positive liquidity premium, long-term interest rates will be substantially higher than current short-term rates, and the yield curve will have a steep upward slope. Conversely, if short-term rates are high, people usually expect them to come back down. Long-term rates will then drop below short-term rates because the average of expected future short-term rates will be far below current short-term rates, despite positive liquidity premiums, and so the yield curve will slope downward.

The liquidity premium and preferred habitat theories explain fact 3, which states that yield curves typically slope upward, by recognizing that the liquidity premium rises with a bond's maturity because of investors' preferences for short-term bonds. Even if short-term interest rates are expected to stay the same, on average, in the future, long-term interest rates will be above short-term interest rates, and yield curves will typically slope upward.

How can the liquidity premium and preferred habitat theories explain the occasional appearance of inverted yield curves if the liquidity premium is positive? It must be that, at times, short-term interest rates are expected to fall so much in the future that the average of the expected short-term rates is well below the current short-term rate. Even when the positive liquidity premium is added to this average, the resulting long-term rate is still lower than the current short-term interest rate.

As our discussion indicates, a particularly attractive feature of the liquidity premium and preferred habitat theories is that you can see what the market is predicting about future short-term interest rates just from the slope of the yield curve. A steeply rising yield curve, as in panel (a) of Figure 6-6, indicates that short-term interest rates are expected to rise in the future. A moderately steep yield curve, as in panel (b), indicates that short-term interest rates are not expected to rise or fall much in the future.

Mini-lecture

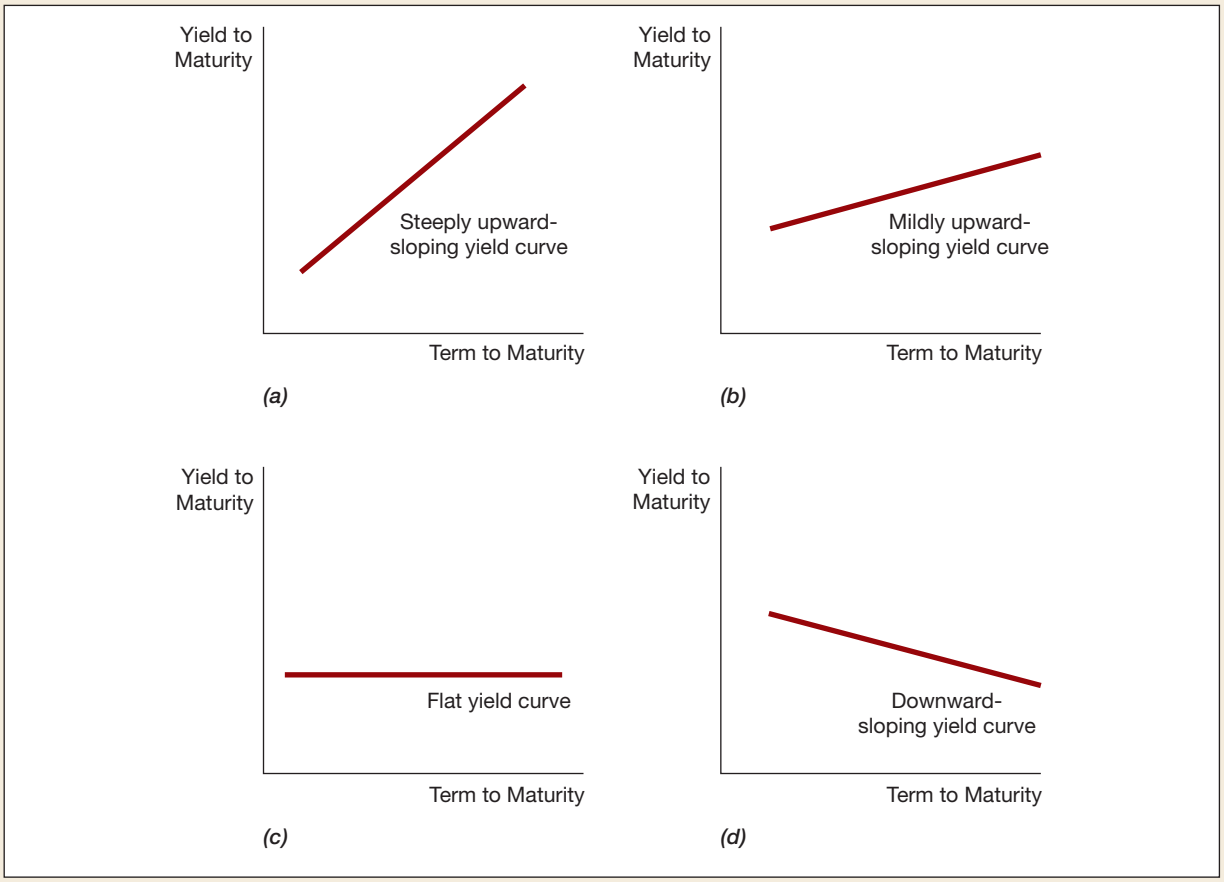


FIGURE 6-6 Yield Curves and the Market's Expectations of Future Short-Term Interest Rates According to the Liquidity Premium (Preferred Habitat) Theory

A steeply rising yield curve, as in panel (a), indicates that short-term interest rates are expected to rise in the future. A moderately steep yield curve, as in panel (b), indicates that short-term interest rates are not expected to rise or fall much in the future. A flat yield curve, as in panel (c), indicates that short-term rates are expected to fall moderately in the future. Finally, an inverted yield curve, as in panel (d), indicates that short-term interest rates are expected to fall sharply in the future.

FYI The Yield Curve as a Forecasting Tool for Inflation and the Business Cycle

Because the yield curve contains information about future expected interest rates, it should have the capacity to help forecast inflation and real output fluctuations. To see why, recall from Chapter 5 that rising interest rates are associated with economic booms and falling interest rates with recessions. When the yield curve is either flat or downward-sloping, it suggests that future short-term interest rates are expected to fall because the economy is more likely to enter a recession. Indeed, the yield curve is found to be an accurate predictor of the business cycle.

In Chapter 4, we learned that a nominal interest rate is composed of a real interest rate and expected inflation, implying that the yield curve contains information about both the future path of nominal interest rates and future inflation. A steep yield curve predicts a future increase in inflation, while a flat or downward-sloping yield curve forecasts a future decline in inflation.

Because the yield curve can be used to forecast business cycles and inflation, the slope of the yield curve is part of the tool kit of many economic forecasters and is often viewed as a useful indicator of the stance of monetary policy, with a steep yield curve indicating loose policy and a flat or downward-sloping yield curve indicating tight policy.

A flat yield curve, as in panel (c), indicates that short-term rates are expected to fall moderately in the future. Finally, an inverted yield curve, as in panel (d), indicates that short-term interest rates are expected to fall sharply in the future.

Evidence on the Term Structure

In the 1980s, researchers examining the term structure of interest rates questioned whether the slope of the yield curve provides information about movements of future short-term interest rates. They found that the spread between long- and short-term interest rates does not always help predict future short-term interest rates, a finding that may stem from substantial fluctuations in the liquidity (term) premium for long-term bonds. More recent research using more discriminating tests now favours a different view. It shows that the term structure contains quite a bit of information for the very short run (over the next several months) and the long run (over several years) but is unreliable at predicting movements in interest rates over the intermediate term (the time in between). Research also has found that the yield curve can be used to help forecast future inflation and business cycles (see the FYI box).

Summary

The liquidity premium and preferred habitat theories are the most widely accepted theories of the term structure of interest rates because they explain the major empirical facts about the term structure so well. They combine the features of both the expectations theory and the segmented markets theory by asserting that a long-term interest rate will be the sum of a liquidity (term) premium and the average of the short-term interest rates that are expected to occur over the life of the bond.

The liquidity premium and preferred habitat theories explain the following facts: (1) Interest rates on bonds of different maturities tend to move together over time; (2) when short-term interest rates are low, yield curves are more likely to have a steep

upward slope; (3) yield curves usually slope upward, but when short-term interest rates are high, yield curves are more likely to be inverted.

The theories also help us predict the movement of short-term interest rates in the future. A steep upward slope of the yield curve means that short-term rates are expected to rise, a mild upward slope means that short-term rates are expected to remain the same, a flat slope means that short-term rates are expected to fall moderately, and an inverted yield curve means that short-term rates are expected to fall sharply.

APPLICATION Interpreting Yield Curves, 1986–2020

Figure 6-7 illustrates several yield curves that have appeared for Canadian government bonds in recent years. What do these yield curves tell us about the public's expectations of future movements of short-term interest rates?

The inverted yield curve that occurred in January 1990 indicated that short-term interest rates were expected to decline sharply in the future. In order for longer-term interest rates with their positive liquidity premium to be well below the short-term interest rate, short-term interest rates must be expected to decline so sharply that their average is far

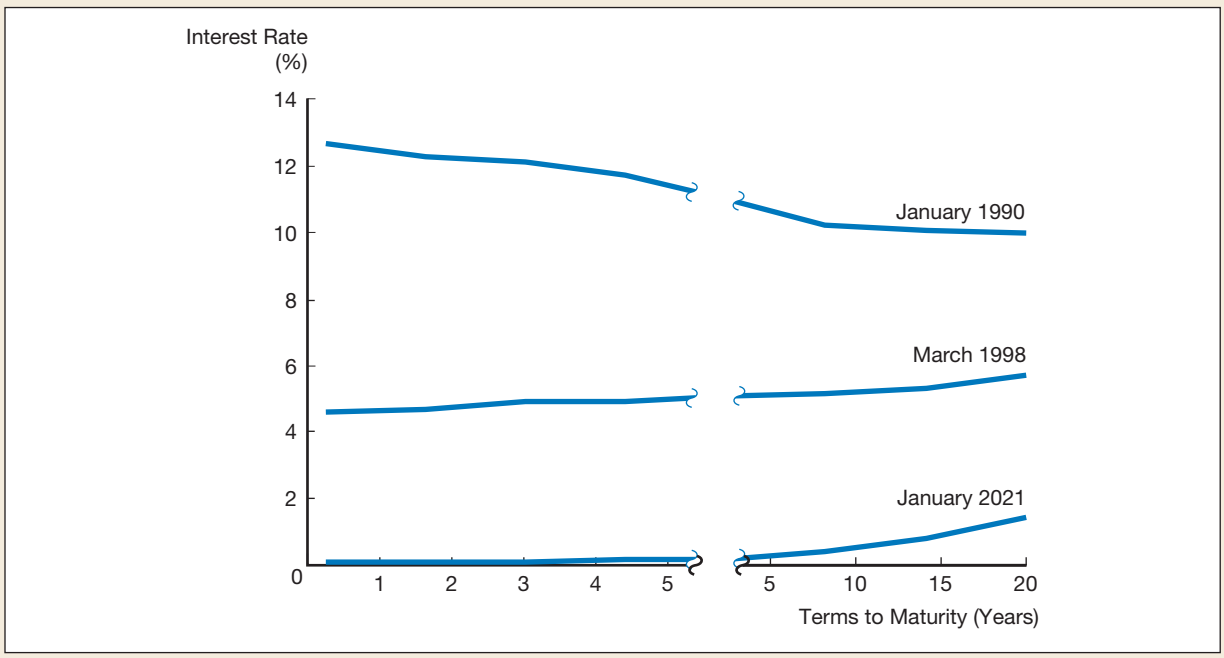


FIGURE 6-7 Yield Curves for Government of Canada Bonds

Yield curves for government of Canada bonds for different dates from 1986 to 2021.

Sources: Statistics Canada CANSIM series V122531, V122532, V122533, V122538, V122539, V122540, V122543, and V122544, and the authors' calculations.

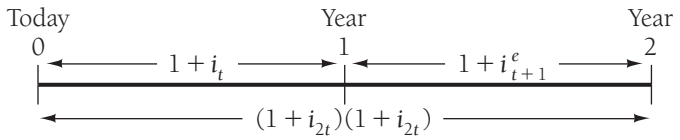
below the current short-term rate. Indeed, the public's expectations of sharply lower short-term interest rates evident in the yield curve were realized soon after January 1990; by March 1991, three-month Treasury bill rates had declined from over 12% to less than 9%.

The upward-sloping yield curve in January 2021 indicates that short-term interest rates would climb in the future. The long-term interest rate is above the short-term interest rate when short-term interest rates are expected to rise because their average plus the liquidity premium will be above the current short-term rate. The moderately upward-sloping yield curve in March 1998 indicates that short-term interest rates were expected neither to rise nor to fall in the near future. In this case, their average remains the same as the current short-term rate, and the positive liquidity premium for longer-term bonds explains the moderate upward slope of the yield curve.

APPLICATION Using the Term Structure to Forecast Interest Rates

Interest-rate forecasts are extremely important to managers of financial institutions because future changes in interest rates have a significant impact on the profitability of their institutions. Furthermore, interest-rate forecasts are needed when managers of financial institutions have to set interest rates on loans that are promised to customers in the future. Our discussion of the term structure of interest rates has indicated that the slope of the yield curve provides general information about the market's prediction of the future path of interest rates. For example, a steeply upward-sloping yield curve indicates that short-term interest rates are predicted to rise in the future, and a downward-sloping yield curve indicates that short-term interest rates are predicted to fall. However, a financial institution manager needs much more specific information on interest-rate forecasts than this. Here we show how the manager of a financial institution can generate specific forecasts of interest rates using the term structure.

To see how this is done, let's start the analysis using the approach we took in developing the expectations theory. Recall that because bonds of different maturities are considered perfect substitutes, we assumed that the expected return over two periods from investing \$1 in a two-period bond, which is $(1 + i_{2t})(1 + i_{2t}) - 1$, must equal the expected return from investing \$1 in one-period bonds, which is $(1 + i_t)(1 + i^e_{t+1}) - 1$. This is shown graphically as follows:



In other words,

$$(1 + i_t)(1 + i^e_{t+1}) - 1 = (1 + i_{2t})(1 + i_{2t})(1 + i_{2t}) - 1$$

Through some tedious algebra we can solve for i^e_{t+1}

$$i^e_{t+1} = \frac{(1 + i_{2t})^2}{1 + i_t} - 1 \tag{4}$$

This measure of i^e_{t+1} is called the forward rate because it is the one-period interest rate that the pure expectations theory of the term structure indicates is expected to prevail one period in the future. To differentiate forward rates derived from the term structure from actual interest rates that are observed at time t, we call these observed interest rates spot rates.

Going back to the expectations theory and the yield curve, which we used to discuss the expectations theory earlier in this chapter, at time t, the one-year interest rate is 5% and the two-year rate is 5.5%. Plugging these numbers into Equation 4 yields the following estimate of the forward rate one period in the future:

$$i^e_{t+1} = \frac{(1 + 0.055)^2}{1 + 0.05} - 1 = 0.06 = 6\%$$

Not surprisingly, this 6% forward rate is identical to the expected one-year interest rate one year in the future. This is exactly what we should find, as our calculation here is just another way of looking at the pure expectations theory.

We can also compare holding the three-year bond against holding a sequence of one-year bonds, which reveals the following relationship:

$$(1 + i_t)(1 + i^e_{t+1})(1 + i^e_{t+2}) - 1 = (1 + i_{3t})(1 + i_{3t})(1 + i_{3t}) - 1$$

And plugging in the estimate for i^e_{t+1} derived in Equation 4, we can solve for i^e_{t+2}

$$i^e_{t+2} = \frac{(1 + i_{3t})^3}{(1 + i_{2t})^2} - 1$$

Continuing with these calculations, we obtain the general solution for the forward rate n periods into the future:

$$i^e_{t+n} = \frac{(1 + i_{n+1t})^{n+1}}{(1 + i_{nt})^n} - 1 \tag{5}$$

Our discussion indicated that the expectations theory is not entirely satisfactory because investors must be compensated with liquidity premiums to induce them to hold longer-term bonds. Hence, we need to modify our analysis, as we did when discussing the liquidity premium theory, by allowing for these liquidity premiums in estimating predictions of future interest rates.

Recall from the discussion of those theories that because investors prefer to hold short-term rather than long-term bonds, the n-period interest rate differs from that indicated by the pure expectations theory by a liquidity premium of l_{nt}. So to allow for liquidity premiums, we need merely subtract l_{nt} from i_{nt} in our formula to derive i^e_{t+n}:

$$i^e_{t+n} = \frac{(1 + i_{n+1t} - l_{n+1t})^{n+1}}{(1 + i_{nt} - l_{nt})^n} - 1 \tag{6}$$

This measure of i^e_{t+n} is referred to, naturally enough, as the *adjusted forward-rate forecast*.

In the case of i^e_{t+1}, Equation 6 produces the following estimate:

$$i^e_{t+1} = \frac{(1 + i_{2t} - l_{2t})^2}{1 + i_t} - 1$$

According to the liquidity premium theory and the yield curve, at time t the l_{2t} liquidity premium is 0.25%, $l_{1t} = 0$, the one-year interest rate is 5%, and the two-year interest

rate is 5.75%. Plugging these numbers into our equation yields the following adjusted forward-rate forecast for one period in the future:

$$i^e_{t+1} = \frac{(1 + 0.0575 - 0.0025)^2}{1 + 0.05} - 1 = 0.06 = 6\%$$

which is the same as the expected interest rate used in the expectations theory, as it should be.

Our analysis of the term structure thus provides managers of financial institutions with a fairly straightforward procedure for producing interest-rate forecasts. First they need to estimate l_{nt}, the values of the liquidity premiums for various n. Then they need merely apply the formula in Equation 6 to derive the market's forecasts of future interest rates.

APPLICATION Forward Rate

A customer asks a bank if it would be willing to commit to making the customer a one-year loan at an interest rate of 8% one year from now. To compensate for the costs of making the loan, the bank needs to charge one percentage point more than the expected interest rate on a Canada bond with the same maturity if it is to make a profit. If the bank manager estimates the liquidity premium to be 0.4%, and the one-year Canada bond rate is 6% and the two-year bond rate is 7%, should the manager be willing to make the commitment?

Solution
The bank manager is unable to make the loan because at an interest rate of 8%, the loan is likely to be unprofitable to the bank.

$$i^e_{t+n} = \frac{(1 + i_{n+1t} - l_{n+1t})^{n+1}}{(1 + i_{nt} - l_{nt})^n} - 1$$

where

$$\begin{aligned} i_{n+1t} &= \text{two-year bond rate} = 0.07 \\ l_{n+1t} &= \text{liquidity premium} = 0.004 \\ i_{nt} &= \text{one-year bond rate} = 0.06 \\ l_{1t} &= \text{liquidity premium} = 0 \\ n &= \text{number of years} = 1 \end{aligned}$$

Thus

$$i^e_{t+1} = \frac{(1 + 0.07 - 0.004)^2}{1 + 0.06} - 1 = 0.072 = 7.2\%$$

The market's forecast of the one-year Canada bond rate one year in the future is therefore 7.2%. Adding the 1% necessary to make a profit on the one-year loan means that the loan is expected to be profitable only if it has an interest rate of 8.2% or higher. ◆

SUMMARY

1. Bonds with the same maturity will have different interest rates because of three factors: default risk, liquidity, and tax considerations. The greater a bond's default risk, the higher its interest rate relative to the interest rates of other bonds; the greater a bond's liquidity, the lower its interest rate; and bonds with tax-exempt status will have lower interest rates than they otherwise would. The relationship among interest rates on bonds with the same maturity that arises because of these three factors is known as the *risk structure of interest rates*.
2. Three theories of the term structure provide explanations of how interest rates on bonds with different terms to maturity are related. The expectations theory views long-term interest rates as equalling the average of future short-term interest rates expected to occur over the life of the bond. By contrast, the segmented markets theory treats the determination of interest rates for each bond's maturity as the outcome of supply and demand in that market only. Neither of these theories by itself can explain the fact that interest rates on bonds of different maturities move together over time and that yield curves usually slope upward.
3. The liquidity premium (preferred habitat) theory combines the features of the other two theories and by so doing is able to explain the facts just mentioned. The liquidity premium (preferred habitat) theory views long-term interest rates as equalling the average of future short-term interest rates expected to occur over the life of the bond plus a liquidity premium. The liquidity premium (preferred habitat) theory allows us to infer the market's expectations about the movement of future short-term interest rates from the yield curve. A steeply upward-sloping curve indicates that future short-term rates are expected to rise; a mildly upward-sloping curve, that short-term rates are expected to stay the same; a flat curve, that short-term rates are expected to decline slightly; and an inverted yield curve, that a substantial decline in short-term rates is expected in the future.

MyLab Economics Make the grade with MyLab Economics: Select end-of-chapter questions can be found on MyLab Economics. You can practise them as often as you want, and most feature step-by-step guided instructions to help you find the right answer.

QUESTIONS

1. If junk bonds are "junk," then why do investors buy them?
2. Which should have the higher risk premium on its interest rates, a corporate bond with a Moody's Baa rating or a corporate bond with a C rating? Why?
3. Do you think that a Canadian Treasury bill will have a risk premium that is higher than, lower than, or the same as that of a similar security (in terms of maturity and liquidity) issued by the government of Colombia?
4. In the fall of 2008, AIG, the largest insurance company in the world at the time, was at risk of defaulting due to the severity of the global financial crisis. As a result, the U.S. government stepped in to support AIG with large capital injections and an ownership stake. How would this affect, if at all, the yield and risk premium on AIG corporate debt?
5. Risk premiums on corporate bonds are usually *anticyclical*; that is, they decrease during business cycle expansions and increase during recessions. Why is this so?
6. Just before the collapse of the U.S. subprime mortgage market in 2007, the most important credit-rating agencies rated mortgage-backed securities with Aaa and AAA ratings. Explain how it was possible that a few months into 2008, the same securities had the lowest possible ratings. Should we always trust credit-rating agencies?
7. The Canadian government offers some of its debt as Real Return Bonds, in which the price of bonds is adjusted for inflation over the life of the debt instrument. Real Return Bonds are traded on a much smaller scale than nominal Canada bonds of equivalent maturity.

What can you conclude about the liquidity premiums of Real Return Bonds and nominal Canada bonds?

8. Predict what will happen to interest rates on a corporation's bonds if the federal government guarantees today that it will pay creditors if the corporation goes bankrupt in the future. What will happen to the interest rates on Canada bonds?

9. Predict what would happen to the risk premiums on corporate bonds if brokerage commissions were lower in the corporate bond market.

10. During 2008, the difference in yield (the *yield spread*) between three-month AA-rated financial commercial paper and three-month AA-rated nonfinancial commercial paper steadily increased from its usual level of close to zero, spiking to over a full percentage point at its peak in October 2008. What explains this sudden increase?

11. If the income tax exemption on U.S. municipal bonds were abolished, what would happen to the interest rates on these bonds? What effect would the change have on interest rates on U.S. Treasury securities?

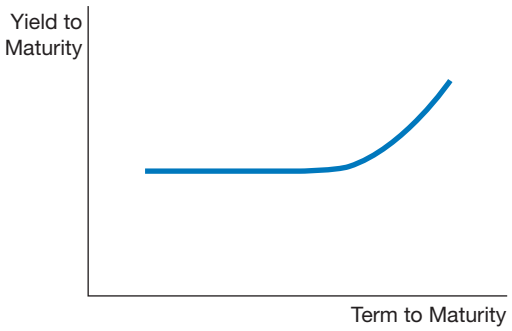
12. Prior to 2008, mortgage lenders in the United States required a house inspection to assess a home's value and often used the same one or two inspection companies in the same geographic market. Following the collapse of the housing market in 2008, mortgage lenders required a house inspection, but this inspection was arranged through a third party. How does the pre-2008 scenario illustrate a conflict of interest similar to the role that credit-rating agencies played in the global financial crisis?

13. "According to the expectations theory of the term structure, it is better to invest in one-year bonds, reinvested over two years, than to invest in a two-year bond if interest rates on one-year bonds are expected to be the same in both years." Is this statement true, false, or uncertain?

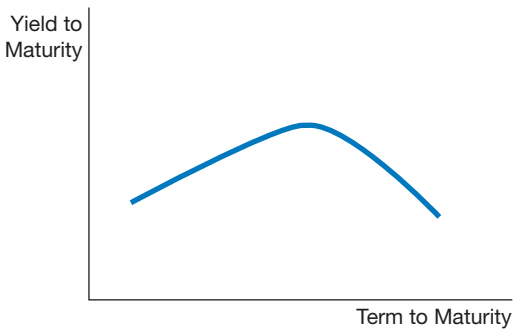
14. If bond investors decide that 30-year bonds are no longer as desirable an investment as they were previously, predict what will happen to the yield curve, assuming (a) the expectations theory of the term structure holds; and (b) the segmented markets theory of the term structure holds.

15. Suppose the interest rates on 1-, 5-, and 10-year Canada bonds are currently 3%, 6%, and 6%, respectively. Investor A chooses to hold only 1-year bonds, and Investor B is indifferent with regard to holding 5- and 10-year bonds. How can you explain the behaviour of Investors A and B?

16. If a yield curve looks like the one shown in the figure below, what is the market predicting about the movement of future short-term interest rates? What might the yield curve indicate about the market's predictions for the inflation rate in the future?



17. If a yield curve looks like the one shown in the figure below, what is the market predicting about the movement of future short-term interest rates? What might the yield curve indicate about the market's predictions for the inflation rate in the future?



18. If yield curves, on average, were flat, what would this say about the liquidity (term) premiums in the term structure? Would you be more or less willing to accept the expectations theory?

19. If the yield curve suddenly became steeper, how would you revise your predictions of interest rates in the future?

20. If expectations of future short-term interest rates suddenly fell, what would happen to the slope of the yield curve?

21. Suppose that the Bank of Canada made an announcement that it would purchase up to $300 billion of longer-term Canada securities over the following six months. What effect might this policy have on the yield curve?

APPLIED PROBLEMS

22. In 2010 and 2011, the government of Greece risked defaulting on its debt due to a severe budget crisis. Using bond market graphs, compare the effects on the risk premium between U.S. Treasury debt and comparable-maturity Greek debt.

23. Assuming the expectations theory is the correct theory of the term structure, calculate the interest rates in the term structure for maturities of one to five years, and plot the resulting yield curves for the following paths of one-year interest rates over the next five years:

a. 5%, 7%, 7%, 7%, 7%

b. 5%, 4%, 4%, 4%, 4%

How would your yield curves change if people preferred shorter-term bonds to longer-term bonds?

24. Assuming the expectations theory is the correct theory of the term structure, calculate the interest rates in the term structure for maturities of one to five years, and plot the resulting yield curves for the following paths of one-year interest rates over the next five years:

a. 5%, 6%, 7%, 6%, 5%

b. 5%, 4%, 3%, 4%, 5%

How would your yield curves change if people preferred shorter-term bonds over longer-term bonds?

25. The table below shows current and expected future one-year interest rates as well as current interest rates on multiyear bonds. Use the table to calculate the liquidity premium for each multiyear bond.

Year	One-Year Bond Rate	Multiyear Bond Rate
1	2%	2%
2	3%	3%
3	4%	5%
4	6%	6%
5	7%	8%

DATA ANALYSIS PROBLEMS

1. Go to the CANSIM database and download monthly data, from January 1976 to January 2021, on the three-month T-bill rate (series V122531) and the interest rate on long-term Canada bonds (series V122544).

a. Calculate the mean and standard deviation as well as the maximum and minimum values for each series over the sample period.

b. When is the spread between the long-term rate and the T-bill rate the highest? The lowest?

2. Go to the CANSIM database and find monthly yield data from January 1983 to January 2021 on the following government securities: one-month (series V122529), three-month (series V122531), six-month (series V122532), 1-year (series V122533), 2-year (series V122538), 3-year (series V122539), 5-year (series V122540), 10-year (series V122543), and long-term (series V122544). Download the last full year of data available into a spreadsheet.

a. Construct a yield curve by creating a line graph for January 2021 and for the same month in 2020, across all the maturities.

b. How do the yield curves compare? What does the changing slope say about potential changes in economic conditions?

3. **Real-Time Data Analysis** Go to the St. Louis Federal Reserve FRED database and find data on Moody's Aaa corporate bond yield (AAA) and Moody's Baa corporate bond yield (BAA). Download the data into a spreadsheet.

a. Calculate the spread (difference) between the Baa and Aaa corporate bond yields for the most recent month of data available. What does this difference represent?

b. Calculate the spread again, for the same month but one year prior, and compare the result to your answer to part (a). What do your answers say about how the risk premium has changed over the past year?

c. Identify the month of highest and lowest spreads since the beginning of the year 2000. How do these spreads compare to the most current spread data available? Interpret the results.

7 The Stock Market, the Theory of Rational Expectations, and the Efficient Market Hypothesis

Learning Objectives

7.1 Calculate the price of common stock.

7.2 Recognize the impact of new information on stock prices.

7.3 Compare and contrast adaptive expectations and rational expectations.

7.4 Explain why arbitrage opportunities imply that the efficient market hypothesis holds.

7.5 Identify and explain the implications of the efficient market hypothesis for financial markets.

7.6 Summarize the reasons why behavioural finance suggests that the efficient market hypothesis may not hold.

Preview

Rarely does a day go by that the stock market isn't a major news item. We have witnessed huge swings in the stock market in recent years. The 1990s were an extraordinary decade for stocks: the S&P/TSX Composite in Canada and the Dow Jones and S&P 500 indexes in the United States increased more than 400%, while the tech-laden NASDAQ index rose more than 1000%. By early 2000, these indexes had reached record highs. Unfortunately, the good times did not last. Starting in early 2000, the stock market began to decline and many investors lost their shirts. The S&P/TSX crashed, falling by more than 40% through March 2003. After subsequently rising over 120%, the stock market crashed again during the global financial crisis, falling by over 40% from its peak in the fall of 2007. Starting in 2009, the stock market recovered quickly, more than quadrupling by February 2020, only to crash with the onset of the coronavirus pandemic.

Because so many people invest in the stock market and the prices of stocks affect the ability of people to retire comfortably, the market for stocks is undoubtedly the financial market that receives the most attention and scrutiny. In this chapter, we look first at how this important market works.

We begin by discussing the fundamental theories that underlie the valuation of stocks. These theories are critical to understanding the forces that cause the value of stocks to rise and fall minute by minute and day by day. Once we have learned the methods for stock valuation, we need to explore how expectations about the market affect its behaviour. We do so by examining the *theory of rational expectations*. When this theory is applied to financial markets, the outcome is the *efficient market hypothesis*, which has some general implications for how markets in other securities besides stocks operate. The theory of rational expectations is also central to debates about the conduct of monetary policy, to be discussed in Chapter 25.

7.1 COMPUTING THE PRICE OF COMMON STOCK

LO 7.1 Calculate the price of common stock.

Common stock is the principal medium through which corporations raise equity capital. **Stockholders**—those who hold stock in a corporation—own an interest in the corporation equal to the percentage of outstanding shares they own. This ownership interest gives them a bundle of rights. The most important are the right to vote and to be a

residual claimant of all funds flowing into the firm (known as **cash flows**), meaning that the stockholder receives whatever remains after all other claims against the firm's assets have been satisfied. Stockholders may receive dividends from the net earnings of the corporation. **Dividends** are payments made periodically, usually every quarter, to stockholders. The board of directors of the firm sets the level of the dividend, usually based on the recommendation of management. In addition, the stockholder has the right to sell the stock.

One basic principle of finance is that the value of any investment is calculated by computing the present value of all cash flows the investment will generate over its life. For example, a commercial building will sell for a price that reflects the net cash flows (rents minus expenses) it is projected to have over its useful life. Similarly, we value common stock as the value in today's dollars of all future cash flows. The cash flows that a stockholder might earn from stock are dividends, the sales price, or both.

To develop the theory of stock valuation, we begin with the simplest possible scenario: You buy the stock, hold it for one period to get a dividend, then sell the stock. We call this the *one-period valuation model*.

The One-Period Valuation Model

Suppose that you have some extra money to invest for one year. After a year, you will need to sell your investment to pay tuition. After watching a report on BNN or Bloomberg, you decide that you want to buy Intel Corp. stock. You call your broker and find that Intel is currently selling for $50 per share and pays $0.16 per year in dividends. The analyst on the financial news predicts that the stock will be selling for $60 in one year. Should you buy this stock?

To answer this question, you need to determine whether the current price accurately reflects the analyst's forecast. To value the stock today, you need to find the present discounted value of the expected cash flows (future payments) using the formula in Equation 1 of Chapter 4. In this equation, the discount factor used to discount the cash flows is the required return on investments in equity rather than the interest rate. The cash flows consist of one dividend payment plus a final sales price. When these cash flows are discounted back to the present, the following equation computes the current price of the stock:

$$P_0 = \frac{D_1}{(1 + k_e)} + \frac{P_1}{(1 + k_e)} \tag{1}$$

where
- P_0 = the current price of the stock. The zero subscript refers to time period zero, or the present.
- D_1 = the dividend paid at the end of year 1
- k_e = the required return on investments in equity
- P_1 = the price at the end of year 1; the predicted sales price of the stock

To see how Equation 1 works, let's compute the price of the Intel stock if, after careful consideration, you decide that you would be satisfied to earn a 12% return on the investment. If you have decided that $k_e = 0.12$, are told that Intel pays $0.16 per year in dividends $(D_1 = 0.16)$, and forecast the share price of $60 at the end of year 1 $(P_1 = \$60)$, you get the following result from Equation 1:

$$P_0 = \frac{0.16}{1 + 0.12} + \frac{\$60}{1 + 0.12} = \$0.14 + \$53.57 = \$53.71$$

On the basis of your analysis, you find that the present value of all cash flows from the stock is $53.71. Because the stock is currently priced at $50 per share, you would choose to buy it. However, you should be aware that the stock may be selling for less than $53.71 because other investors have placed a greater risk on the cash flows or estimated the cash flows to be less than you did.

The Generalized Dividend Valuation Model

Using the present value concept, we can extend the one-period dividend valuation model to any number of periods: The value of a stock today is the present value of all future cash flows. The only cash flows that an investor will receive are dividends and a final sales price when the stock is ultimately sold in period n. The generalized multiperiod formula for stock valuation can be written as

$$P_0 = \frac{D_1}{(1 + k_e)^1} + \frac{D_2}{(1 + k_e)^2} + \ \ldots \ + \frac{D_n}{(1 + k_e)^n} + \frac{P_n}{(1 + k_e)^n} \tag{2}$$

where P_n = the price of the stock at the end of period n
D_i = the dividend paid at the end of year i

If you tried to use Equation 2 to find the value of a share of stock, you would soon realize that you must first estimate the value the stock will have at some point in the future before you can estimate its value today. In other words, you must find P_n before you can find P_0. However, if P_n is far in the future, it will not affect P_0. For example, the present value of a share of stock that sells for $50 seventy-five years from now, using a 12% discount rate, is just one cent $[\$50/(1.12^{75}) = \$0.01]$. This reasoning implies that the current value of a share of stock can be calculated as simply the present value of the future dividend stream. The **generalized dividend model** is given in Equation 3. Note that it is the same formula as Equation 2 but without the final sales price:

$$P_0 = \sum_{t=1} \frac{D_t}{(1 + k_e)^t} \tag{3}$$

Consider the implications of Equation 3 for a moment. The generalized dividend model states that the price of a stock is determined only by the present value of the dividends and that nothing else matters. Many stocks do not pay dividends, so how is it that these stocks have value? *Buyers of the stock expect that the firm will pay dividends someday*. Most of the time a firm institutes dividends as soon as it has completed the rapid growth phase of its life cycle.

The generalized dividend valuation model requires that we compute the present value of an infinite stream of dividends, a process that could be difficult, to say the least. Therefore, simplified models have been developed to make the calculations easier. One such model is the **Gordon growth model**, which assumes constant dividend growth.

The Gordon Growth Model

Many firms strive to increase their dividends at a constant rate each year. Equation 4 is derived from Equation 3 to reflect this constant growth in dividends:

$$P_0 = \frac{D_0 \times (1 + g)^1}{(1 + k_e)^1} + \frac{D_0 \times (1 + g)^2}{(1 + k_e)^2} + \ \ldots \ + \frac{D_0 \times (1 + g)^\infty}{(1 + k_e)^\infty} \tag{4}$$

where D_0 = the most recent dividend paid
g = the expected constant growth rate in dividends
k_e = the required return on an investment in equity

Equation 4 can be simplified to obtain Equation 5:[1]

$$P_0 = \frac{D_0 \times (1 + g)}{(k_e - g)} = \frac{D_1}{(k_e - g)} \tag{5}$$

This model is useful for finding the value of a stock, given a few assumptions:

1. *Dividends are assumed to continue growing at a constant rate forever.* Actually, as long as the dividends are expected to grow at a constant rate for an extended period of time, the model should yield reasonable results. This is because errors about distant cash flows become small when discounted to the present.
2. *The growth rate is assumed to be less than the required return on equity, k_e.* Myron Gordon, in his development of the model, demonstrated that this is a reasonable assumption. In theory, if the growth rate were faster than the rate demanded by holders of the firm's equity, then in the long run the firm would grow impossibly large.

7.2 HOW THE MARKET SETS STOCK PRICES

LO 7.2 Recognize the impact of new information on stock prices.

Suppose you go to an auto auction. The cars are available for inspection before the auction begins, and you find a little pre-owned Tesla Model 3 that you like. You test-drive it in the parking lot and notice that it makes a few strange noises, but you decide that you still like the car. You decide that $20 000 is a fair price that will allow you to pay some repair bills, should the noises turn out to be serious. You see that the auction is ready to begin, so you go in and wait for the pre-owned Tesla Model 3 to be auctioned.

Suppose another buyer also spots the pre-owned Tesla Model 3. They test-drive the car and recognize that the noises are simply the result of worn brake pads that they can fix themselves at a nominal cost. They decide that the car is worth $22 000. They also go in and wait for the pre-owned Tesla Model 3 to come up for auction.

[1]To generate Equation 5 from Equation 4, first multiply both sides of Equation 4 by $(1 + k_e)/(1 + g)$ and then subtract Equation 4 from the result. This yields

$$\frac{P_0 \times (1 + k_e)}{(1 + g)} - P_0 = D_0 - \frac{D_0 \times (1 + g)^{\infty}}{(1 + k_e)^{\infty}}$$

Assuming that k_e is greater than g, the term on the far right will approach zero and so can be dropped. Thus, after factoring P_0 out of the left-hand side,

$$P_0 \times \left[\frac{1 + k_e}{1 + g} - 1\right] = D_0$$

Next, simplify by combining terms:

$$P_0 \times \frac{(1 + k_e) - (1 + g)}{1 + g} = D_0$$

$$P_0 = \frac{D_0 \times (1 + g)}{k_e - g} = \frac{D_1}{k_e - g}$$

Who will buy the car, and for how much? Suppose only the two of you are interested in the pre-owned Tesla Model 3. You begin the bidding at $18 000. Your competitor ups your bid to $19 000. You bid your top price of $20 000. They counter with $21 000. The price is now higher than you are willing to pay, so you stop bidding. The car is sold to the more informed buyer for $21 000.

This simple example raises a number of important points. First, the price is set by the buyer who is willing to pay the highest price. This price is not necessarily the highest price the asset could fetch, but it is incrementally greater than what any other buyer is willing to pay.

Second, the market price will be set by the buyer who can take best advantage of the asset. The buyer who purchased the car knew that they could fix the noise easily and cheaply. As a consequence, they were willing to pay more for the car than you were. The same concept holds for other assets. For example, a piece of property or a building will sell to the buyer who can put the asset to the most productive use.

Finally, the example shows the role played by information in asset pricing. Superior information about an asset can increase its value by reducing its risk. When you consider buying a stock, the future cash flows are subject to many unknowns. The buyer who has the best information about these cash flows will discount them at a lower interest rate than will a buyer who is very uncertain.

Now let's apply these ideas to stock valuation. Suppose you are considering the purchase of stock expected to pay a $2 dividend next year. Market analysts expect the firm to grow at 3% indefinitely. You are *uncertain* about both the constancy of the dividend stream and the accuracy of the estimated growth rate. To compensate yourself for this uncertainty (risk), you require a return of 15%.

Now suppose Jennifer, another investor, has spoken with industry insiders and feels more confident about the projected cash flows. Jennifer requires only a 12% return because her perceived risk is lower than yours. Bud, however, is dating the CEO of the company. He knows with more certainty what the future of the firm actually looks like, and thus requires only a 10% return.

What value will each investor give to the stock? Applying the Gordon growth model yields the following stock prices:

Investor	Discount Rate	Stock Price
You	15%	$16.67
Jennifer	12%	$22.22
Bud	10%	$28.57

You are willing to pay $16.67 for the stock. Jennifer will pay up to $22.22, and Bud will pay $28.57. The investor with the lowest perceived risk is willing to pay the most for the stock. If there were no other traders but these three, the market price would be between $22.22 and $28.57. If you already held the stock, you would sell it to Bud.

We thus see that the players in the market, bidding against one another, establish the market price. When new information is released about a firm, expectations change, and with them, prices change. New information can cause changes in expectations about the level of future dividends or the risk of those dividends. Because market participants are constantly receiving new information and revising their expectations, it is reasonable that stock prices are constantly changing as well.

APPLICATION Monetary Policy and Stock Prices

Stock market analysts tend to hang on every word uttered by the governor of the Bank of Canada because they know that an important determinant of stock prices is monetary policy. But how does monetary policy affect stock prices?

The Gordon growth model in Equation 5 explains this relationship. Monetary policy can affect stock prices in two ways. First, when the Bank of Canada lowers interest rates, the return on bonds (an alternative asset to stocks) declines, and investors are likely to accept a lower required rate of return on an investment in equity (k_e). The resulting decline in k_e lowers the denominator in the Gordon growth model (Equation 5), leads to a higher value of P_0, and raises stock prices. Furthermore, a lowering of interest rates is likely to stimulate the economy, so the growth rate in dividends, g, is likely to be somewhat higher. This rise in g also causes the denominator in Equation 5 to decrease, which also leads to a higher P_0 and a rise in stock prices.

As we will see in Chapter 26, the impact of monetary policy on stock prices is one of the key ways in which monetary policy affects the economy. ◆

APPLICATION The Coronavirus Stock Market Crash of 2020

The spread of the coronavirus in February 2020 triggered a stock market crash in which the Dow Jones Industrial Average fell from a peak of 29 551 on February 12 to 18 561 on March 20, a decline of 37%. In Canada, the S&P/TSX fell from a peak of 17 944 on February 20 to 11 228 on March 23, also a decline of 37%. Our analysis of stock price valuation, again using the Gordon growth model, can help us understand how the coronavirus pandemic affected stock prices.

The coronavirus pandemic led to a major lockdown of the economy, resulting in a sharp downward revision of the growth prospects for companies, thus lowering the dividend growth rate (g) in the Gordon model. In terms of Equation 5, the resulting increase in the numerator would lead to a decline in P_0 and hence a decline in stock prices.

The coronavirus pandemic also increased uncertainty about the economy and led to a widening of credit spreads. As a result, the required return on investment in equity rose. In terms of Equation 5, a higher k_e produces an increase in the denominator, a decline in P_0, and a general fall in stock prices.

The combination of expectations of a lower dividend growth rate (g) and a higher required return on investment in equity (k_e) led to one of the steepest stock market declines in U.S. history. ◆

7.3 THE THEORY OF RATIONAL EXPECTATIONS

LO 7.3 Compare and contrast adaptive expectations and rational expectations.

The analysis of stock price evaluation we outlined in the previous section depends on people's expectations—especially expectations of cash flows. Indeed, it is difficult to think of any sector of the economy in which expectations are not crucial; this is why it is important to examine how expectations are formed. We do so by outlining the *theory of rational expectations*, currently the most widely used theory to describe the formation of business and consumer expectations.

In the 1950s and 1960s, economists regularly viewed expectations as formed from past experience only. Expectations of inflation, for example, were typically viewed as being an average of past inflation rates. This view of expectation formation, called **adaptive expectations**, suggests that changes in expectations will occur slowly over time, as data for a variable evolve.[2] So, if inflation had formerly been steady at a 5% rate, expectations of future inflation would be 5% too. If inflation rose to a steady rate of 10%, expectations of future inflation would rise toward 10%, but slowly: In the first year, expected inflation might rise only to 6%; in the second year, to 7%; and so on.

The adaptive expectations hypothesis has been faulted on the grounds that people use more information than just past data on a single variable to form their expectations of that variable. Their expectations of inflation will almost surely be affected by their predictions of future monetary policy as well as by current and past monetary policy. In addition, people often change their expectations quickly in the light of new information. To address these objections to the validity of adaptive expectations, John Muth developed an alternative theory of expectations, called **rational expectations**, which can be stated as follows: ***Expectations will be identical to optimal forecasts (the best guess of the future) using all available information.***[3]

What exactly does this mean? To explain it more clearly, let's use the theory of rational expectations to examine how expectations are formed in a situation that most of us will encounter at some point in our lifetime: our drive to work. Suppose that if Joe Commuter travels when it is not rush hour, his trip takes an average of 30 minutes. Sometimes his trip takes 35 minutes; other times, 25 minutes; but the average, non-rush-hour driving time is 30 minutes. If, however, Joe leaves for work during the rush hour, it takes him, on average, an additional 10 minutes to get to work. Given that he leaves for work during the rush hour, the best guess of his driving time—the **optimal forecast**—is 40 minutes.

If the only information available to Joe before he leaves for work related to his driving time is that he is leaving during the rush hour, what does rational expectations theory allow you to predict about Joe's expectations of his driving time? Since the best guess of his driving time, using all available information, is 40 minutes, Joe's expectation should be the same. Clearly, an expectation of 35 minutes would not be rational, because it is not equal to the optimal forecast, or the best guess of the driving time.

Suppose that the next day, given the same conditions and expectations, it takes Joe 45 minutes to drive to work because he hits an abnormally large number of red lights. The day after that, Joe hits all the lights right and it takes him only 35 minutes to drive to work. Do these variations mean that Joe's 40-minute expectation is irrational? No; an expectation of 40 minutes' driving time is still a rational expectation. In both cases, the forecast is off by 5 minutes, so the expectation has not been perfectly accurate each

[2]More specifically, adaptive expectations—say, of inflation—are written as a weighted average of past inflation rates:

$$\pi_t^e = (1 - \lambda)\sum_{j=0}^{\infty} \lambda^j \pi_{t-j}$$

where π_t^e = adaptive expectation of inflation at time t
π_{t-j} = inflation at time $t - j$
λ = a constant between the values of 0 and 1

[3]John Muth, "Rational Expectations and the Theory of Price Movements," *Econometrica* 29 (1961): 315–335.

time. However, the forecast does not have to be perfectly accurate to be rational—it need only be the *best possible* forecast given the available information; that is, it has to be correct *on average*, and the 40-minute expectation meets this requirement. As there is bound to be some randomness in Joe's driving time regardless of driving conditions, an optimal forecast will never be completely accurate.

The example makes the following important point about rational expectations: ***Even though a rational expectation equals the optimal forecast using all available information, a prediction based on it may not always be perfectly accurate.***

What if an item of information relevant to predicting driving time is unavailable or ignored? Suppose that on Joe's usual route to work, an accident occurs and causes a two-hour traffic jam. If Joe has no way of ascertaining this information, his rush-hour expectation of 40 minutes' driving time is still rational, because the accident information is not available to him for incorporation into his optimal forecast. However, if there was a radio or TV traffic report about the accident that Joe did not bother to listen to, or heard but ignored, his 40-minute expectation is no longer rational. In light of the availability of this information, Joe's optimal forecast should have been 2 hours and 40 minutes.

Accordingly, an expectation may fail to be rational for two reasons:

1. People might be aware of all available information but find it takes too much effort to make their expectation the best guess possible.
2. People might be unaware of some available relevant information, so their best guess of the future will not be accurate.

Nonetheless, it is important to recognize that if an additional factor is important but information about it is not available, an expectation that does not take that factor into account can still be rational.

Formal Statement of the Theory

We can state the theory of rational expectations somewhat more formally. If X stands for the variable that is being forecast (in our example, Joe Commuter's driving time), X^e for the expectation of this variable (Joe's expectation of his driving time), and X^{of} for the optimal forecast of X using all available information (the best guess possible of Joe's driving time), the theory of rational expectations then simply states

$$X^e = X^{of} \tag{6}$$

That is, the expectation of X equals the optimal forecast using all available information.

Rationale Behind the Theory

Why do people try to make their expectations match their best possible guess of future results, using all available information? The simplest explanation is that it is costly for people not to do so. Joe Commuter has a strong incentive to make his expectation of the time it takes him to drive to work as accurate as possible. If he underpredicts his driving time, he will often be late to work and risk being fired. If he overpredicts, he will, on average, get to work too early and will have given up sleep or leisure time unnecessarily. Accurate expectations are desirable, and therefore people have a strong incentive to try to make expectations equal to optimal forecasts by using all available information.

The same principle applies to businesses. Suppose an appliance manufacturer—say, General Electric—knows that interest-rate movements are important to the sales of appliances. If GE makes poor forecasts of interest rates, it will earn less profit, because it will produce either too many or too few appliances. The incentives are strong for GE to acquire all available information to help it forecast interest rates and make the best possible guess of future interest-rate movements.

Implications of the Theory

Rational expectations theory leads to two commonsense implications regarding the formation of expectations. These implications are important in the analysis of both the stock market and the aggregate economy:

1. ***If there is a change in the way a variable moves, the way in which expectations of this variable are formed will change as well.*** This tenet of rational expectations theory can be most easily understood through a concrete example. Suppose interest rates move in such a way that they tend to return to a "normal" level over time. If today's interest rate is high relative to the normal level, an optimal forecast of the interest rate in the future is that it will decline to the normal level. Rational expectations theory would imply that when today's interest rate is high, the expectation is that it will fall in the future.

   Suppose now that the way in which the interest rate moves changes so that when the interest rate is high, it stays high. In this case, when today's interest rate is high, the optimal forecast of the future interest rate, and hence the rational expectation, is that it will stay high. Expectations of the future interest rate will no longer indicate that the interest rate will fall. The change in the way the interest-rate variable moves has therefore led to a change in the way that expectations of future interest rates are formed. The rational expectations analysis here is generalizable to expectations of any variable. Hence, when a change occurs in the way any variable moves, the way in which expectations of this variable are formed will change too.
2. ***The forecast errors of expectations will, on average, be zero and cannot be predicted ahead of time.*** The forecast error of an expectation is $X - X^e$, the difference between the realization of a variable X and the expectation of the variable. That is, if Joe Commuter's driving time on a particular day is 45 minutes and his expectation of the driving time is 40 minutes, the forecast error is 5 minutes.

   Suppose that in violation of the rational expectations tenet, Joe's forecast error is not, on average, equal to zero; instead, it equals 5 minutes. The forecast error is now predictable ahead of time because Joe will soon notice that he is, on average, 5 minutes late for work and can improve his forecast by increasing it by 5 minutes. Rational expectations theory implies that this is exactly what Joe will do because he will want his forecast to be the best guess possible. When Joe has revised his forecast upward by 5 minutes, on average, the forecast error will equal zero, so it cannot be predicted ahead of time. Rational expectations theory implies that forecast errors of expectations cannot be predicted.

The incentives for equating expectations with optimal forecasts are especially strong in financial markets. In these markets, people with better forecasts of the future get rich. The application of the theory of rational expectations to financial markets (where it is called the **efficient market hypothesis** or the **theory of efficient capital markets**) is thus particularly useful.

7.4 THE EFFICIENT MARKET HYPOTHESIS: RATIONAL EXPECTATIONS IN FINANCIAL MARKETS

LO 7.4 Explain why arbitrage opportunities imply that the efficient market hypothesis holds.

While monetary economists were developing the theory of rational expectations, financial economists were developing a parallel theory of expectations formation for financial markets. It led them to the same conclusion as that of the rational expectations theorists: Expectations in financial markets are equal to optimal forecasts using all available information.[4] Although financial economists, such as Eugene Fama, later a winner of the Nobel Prize in economics, gave their theory another name, calling it the *efficient market hypothesis,* in fact their theory is just an application of rational expectations to the pricing of stocks and other securities.

The efficient market hypothesis is based on the assumption that prices of securities in financial markets fully reflect all available information. You may recall from Chapter 4 that the rate of return from holding a security equals the sum of the capital gain on the security (the change in the price), plus any cash payments, divided by the initial purchase price of the security:

$$R = \frac{P_{t+1} - P_t + C}{P_t} \tag{7}$$

where R = rate of return on the security held from time t to time $t + 1$ (say, the end of 2021 to the end of 2022)
P_{t+1} = price of the security at time $t + 1$, the end of the holding period
P_t = price of the security at time t, the beginning of the holding period
C = cash payment (coupon or dividend payments) made in the period t to $t + 1$

Let's look at the expectation of this return at time t, the beginning of the holding period. Because the current price P_t and the cash payment C are known at the outset, the only variable in the definition of the return that is uncertain is the price next period, P_{t+1}.[5] Denoting expectation of the security's price at the end of the holding period as P^e_{t+1}, the expected return R^e is

$$R^e = \frac{P^e_{t+1} - P_t + C}{P_t}$$

The efficient market hypothesis views expectations of future prices as equal to optimal forecasts using all currently available information. In other words, the market's expectations of future securities prices are rational, so that

$$P^e_{t+1} = P^{of}_{t+1}$$

which in turn implies that the expected return on the security will equal the optimal forecast of the return:

$$R^e = R^{of} \tag{8}$$

[4]The development of the efficient market hypothesis was not wholly independent of the development of rational expectations theory in that financial economists were aware of Muth's work.

[5]There are cases in which C might not be known at the beginning of the period, but that does not make a substantial difference to the analysis. We would in that case assume that not only price expectations but also the expectations of C are optimal forecasts using all available information.

Unfortunately, we cannot observe either R^e or P^e_{t+1}, so the rational expectations equations by themselves do not tell us much about how the financial market behaves. However, if we can devise some way to measure the value of R^e, these equations will have important implications for how prices of securities change in financial markets.

The supply and demand analysis of the bond market that we developed in Chapter 5 showed us that the expected return on a security (the interest rate, in the case of the one-year discount bond) will have a tendency to move toward the equilibrium return that equates the quantity demanded to the quantity supplied. Supply and demand analysis enables us to determine the expected return on a security with the following equilibrium condition: The expected return on a security R^e equals the equilibrium return R^*, which equates the quantity of the security demanded to the quantity supplied; that is,

$$R^e = R^* \tag{9}$$

The academic field of finance explores the factors (risk and liquidity, for example) that influence the equilibrium returns on securities. For our purposes, it is sufficient to know that we can determine the equilibrium return and thus determine the expected return with the equilibrium condition.

We can derive an equation to describe pricing behaviour in an efficient market by using the equilibrium condition to replace R^e with R^* in the rational expectations equation (Equation 8). In this way, we obtain

$$R^{of} = R^* \tag{10}$$

This equation tells us that ***current prices in a financial market will be set so that the optimal forecast of a security's return using all available information equals the security's equilibrium return***. Financial economists state it more simply: In an efficient market, a security's price fully reflects all available information.

Rationale Behind the Hypothesis

To see why the efficient market hypothesis makes sense, we make use of the concept of **arbitrage**, in which market participants (*arbitrageurs*) eliminate **unexploited profit opportunities**, that is, returns on a security that are larger than what is justified by the characteristics of that security. Arbitrage is of two types: *pure arbitrage*, in which the elimination of unexploited profit opportunities involves no risk, and the type of arbitrage we discuss here, in which the arbitrageur takes on some risk when eliminating the unexploited profit opportunities. To see how arbitrage leads to the efficient market hypothesis given a security's risk characteristics, let's look at an example. Suppose the normal return on ExxonMobil common stock is 10% at an annual rate, and its current price P_t is lower than the optimal forecast of tomorrow's price P^{of}_{t+1}, so that the optimal forecast of the return at an annual rate is 50%, which is greater than the equilibrium return of 10%. We are now able to predict that, on average, ExxonMobil's return will be abnormally high, so there is an unexploited profit opportunity. Knowing that, on average, you can earn an abnormally high rate of return on ExxonMobil stock (because $R^{of} > R^*$), you will buy more, which will in turn drive up the stock's current price P_t relative to its expected future price P^{of}_{t+1}, thereby lowering R^{of}. When the current price has risen sufficiently so that R^{of} equals R^* and the efficient market condition (Equation 10) is satisfied, the buying of ExxonMobil stock will stop, and the unexploited profit opportunity will disappear.

Similarly, a security for which the optimal forecast of the return is −5% and the equilibrium return is 10% ($R^{of} < R^*$) would be a poor investment because, on average, the security earns less than the equilibrium return. In such a case, you would sell the security and drive down its current price relative to the expected future price until R^{of} rose to the level of R^* and the efficient market condition was again satisfied. Our discussion can be summarized as follows:

$$R^{of} > R^* \rightarrow P_t \uparrow \rightarrow R^{of} \downarrow$$
$$R^{of} < R^* \rightarrow P_t \downarrow \rightarrow R^{of} \uparrow$$
$$\text{until}$$
$$R^{of} = R^*$$

Another way to state the efficient market condition is this: ***In an efficient market, all unexploited profit opportunities will be eliminated.***

An extremely important factor in this reasoning is that ***not everyone in a financial market must be well informed about a security or have rational expectations for its price to be driven to the point at which the efficient market condition holds***. Financial markets are structured so that many participants can play. As long as a few people (often referred to as the "smart money") keep their eyes open for unexploited profit opportunities, they will eliminate the profit opportunities that appear, because in doing so they make a profit. The efficient market hypothesis makes sense because it does not require that everyone in a market be cognizant of what is happening to every security.

Random-Walk Behaviour of Stock Prices

The term **random walk** describes the movements of a variable whose future values cannot be predicted (are random) because, given today's value, the value of the variable is just as likely to fall as it is to rise. An important implication of the efficient market hypothesis is that stock prices should approximately follow a random walk; that is, ***future changes in stock prices should, for all practical purposes, be unpredictable***. The random-walk implication of the efficient market hypothesis is the one most commonly mentioned in the press because it is the most readily comprehensible to the public. In fact, when people mention the "random-walk theory of stock prices," they are in reality referring to the efficient market hypothesis.

The case for random-walk stock prices can be demonstrated. Suppose people could predict that the price of Happy Feet Corporation (HFC) stock would rise 1% in the coming week. The predicted rate of capital gains and rate of return on HFC stock would then exceed 50% at an annual rate. Since this is very likely to be far higher than the equilibrium rate of return on HFC stock ($R^{of} > R^*$), the efficient market hypothesis indicates that people would immediately buy this stock and bid up its current price. The action would stop only when the predictable change in the price dropped to near zero, so that $R^{of} = R^*$.

Similarly, if people could predict that the price of HFC stock would fall by 1%, the predicted rate of return would be negative ($R^{of} < R^*$), and people would immediately sell the stock. The current price would fall until the predictable change in the price rose back to near zero, where the efficient market condition again would hold. The efficient market hypothesis suggests that the predictable change in stock prices will be near

Global **Should Foreign Exchange Rates Follow a Random Walk?**

Although the efficient market hypothesis is usually applied to the stock market, it can also be used to show that foreign exchange rates, like stock prices, should generally follow a random walk. To see why this is the case, consider what would happen if investors could predict that a currency would rise in value by 1% in the coming week. By buying this currency, they could earn a greater than 50% return at an annual rate, which is likely to be far above the equilibrium return for holding a currency. As a result, investors would immediately buy the currency and bid up its current price, thereby reducing the expected return. The process would only stop when the predictable change in the exchange rate dropped to near zero so that the optimal forecast of the return no longer differed from the equilibrium return. Likewise, if investors could predict that the currency would fall in value by 1% in the coming week, they would sell it until the predictable change in the exchange rate was again near zero. The efficient market hypothesis therefore implies that future changes in exchange rates should, for all practical purposes, be unpredictable; in other words, exchange rates should follow random walks. Indeed, the random-walk behaviour of exchange rates is exactly what is found in the data.

zero, leading to the conclusion that stock prices will generally follow a random walk.[6] As the Global Box "Should Foreign Exchange Rates Follow a Random Walk?" indicates, the efficient market hypothesis suggests that foreign exchange rates should also follow a random walk.

APPLICATION Practical Guide to Investing in the Stock Market

The efficient market hypothesis has numerous applications to the real world. It is especially valuable because it can be applied directly to an issue that concerns many of us: how to get rich (or at least not get poor) by investing in the stock market. A practical guide to investing in the stock market, which we develop here, provides a better understanding of the use and implications of the efficient market hypothesis.

How Valuable Are Reports Published by Investment Advisers?

Suppose you have just read in the financial news that investment advisers are predicting a boom in oil stocks because an oil shortage is developing. Should you proceed to withdraw all of your hard-earned savings from the bank and invest them in oil stocks?

[6]Note that the random-walk behaviour of stock prices is only an *approximation* derived from the efficient market hypothesis. Random-walk behaviour would hold *exactly* only for a stock for which an unchanged price leads to its having the equilibrium return. Then, when the predictable change in the stock price is exactly zero, $R^{of} = R^{*}$.

The efficient market hypothesis tells us that when purchasing a security, we cannot expect to earn an abnormally high return, or a return greater than the equilibrium return. Information in newspapers and in the published reports of investment advisers is readily available to many market participants and is already reflected in market prices. So acting on this information will not yield abnormally high returns, on average. The empirical evidence for the most part confirms that recommendations from investment advisers cannot help us outperform the general market. Indeed, as the FYI box "Should You Hire an Ape as Your Investment Adviser?" suggests, human investment advisers in San Francisco do not, on average, even outperform an orangutan!

Probably no other conclusion is met with more skepticism by students than this one when they first hear it. We all know, or have heard of, someone who has been successful in the stock market for a period of many years. We wonder, "How could someone be so consistently successful if they did not really know how to predict when returns would be abnormally high?" The following story, reported in the press, illustrates why such anecdotal evidence is not reliable.

A get-rich-quick artist invented a clever scam. Every week, they wrote two letters. In letter A, they would pick team A to win a particular football game; in letter B, they would pick the opponent, team B. They would then separate a mailing list into two groups, and they would send letter A to the people in one group and letter B to the people in the other. The following week they would do the same thing, but this time they would send these letters only to the group who had received the first letters containing the correct prediction. After doing this for 10 games, they had a small cluster of people who had received letters predicting the correct winning team for every game. They then mailed a final letter to this group, declaring that since they were obviously an expert predictor of the outcome of football games (they had picked the winning teams 10 weeks in a row), and since their predictions were profitable for the recipients who bet on the games, they would continue to send their predictions only if they were paid a substantial amount of money. When one of their clients figured out what they were up to, the con artist was prosecuted and thrown in jail!

What is the lesson of the story? Even if no forecaster can always accurately predict the market, there always will be some who appear to be consistent winners. A person who has done well regularly in the past cannot guarantee that they will do well in the future. Note that there will also be a group of persistent losers, but you rarely hear about them because no one brags about a poor forecasting record.

Should You Be Skeptical of Hot Tips?

Suppose your broker phones you with a hot tip to buy stock in the Happy Feet Corporation (HFC) because it has just developed a product that is completely effective in curing athlete's foot. The stock price is sure to go up. Should you follow this advice and buy HFC stock?

The efficient market hypothesis indicates that you should be skeptical of such news. If the stock market is efficient, it has already priced HFC stock so that its expected return will equal the equilibrium return. The hot tip is not particularly valuable and will not enable you to earn an abnormally high return.

You might wonder, though, if the hot tip is based on new information and would give you an edge on the rest of the market. If other market participants have gotten this information before you, the answer again is no. As soon as the information hits the street, the unexploited profit opportunity it creates will be quickly eliminated. The stock's price will already reflect the information, and you should expect to realize only

FYI **Should You Hire an Ape as Your Investment Adviser?**

The *San Francisco Chronicle* came up with an amusing way of evaluating how successful investment advisers are at picking stocks. They asked eight analysts to pick five stocks at the beginning of the year and then compared the performance of their stock picks to those chosen by Jolyn, an orangutan living at Marine World/Africa USA in Vallejo, California. Jolyn beat the investment advisers as often as they beat her. Given this result, you might be just as well off hiring an orangutan as your investment adviser as you would be hiring a human being!

the equilibrium return. But if you are one of the first to gain the new information, it *can* do you some good. Only then can you be one of the lucky ones who, on average, will earn an abnormally high return by helping eliminate the profit opportunity by buying HFC stock.

Do Stock Prices Always Rise When There Is Good News?

If you follow the stock market, you might have noticed a puzzling phenomenon: When good news about a corporation, such as a particularly favourable earnings report, is announced, the price of the corporation's stock frequently does not rise. The efficient market hypothesis explains this phenomenon.

Because changes in stock prices are unpredictable, when information is announced that has already been expected by the market, stock prices will remain unchanged. The announcement does not contain any new information that would lead to a change in stock prices. If this were not the case, and the announcement led to a change in stock prices, it would mean that the change was predictable. Because such a scenario is ruled out in an efficient market, ***stock prices will respond to announcements only when the information being announced is new and unexpected***. If the news is expected, no stock price response will occur. This is exactly what the evidence shows: Stock prices do reflect publicly available information.

Sometimes an individual stock price declines when good news is announced. Although this seems somewhat peculiar, it is completely consistent with the workings of an efficient market. Suppose that although the announced news is good, it is not as good as expected. HFC's earnings may have risen by 15%, but if the market expected earnings to rise by 20%, the new information is actually unfavourable, and the stock price declines.

Efficient Market Prescription for the Investor

What does the efficient market hypothesis recommend for investing in the stock market? It tells us that hot tips and investment advisers' published recommendations—all of which make use of publicly available information—cannot help an investor outperform the market. Indeed, it indicates that anyone without better information than other market participants cannot expect to beat the market. So what is an investor to do?

The efficient market hypothesis leads to the conclusion that such an investor (and almost all of us fit into this category) should not try to outguess the market by constantly buying and selling securities. This process does nothing but boost the income of brokers, who earn commissions on each trade.[7] Instead, the investor should pursue a "buy-and-hold" strategy—purchase stocks and hold them for long periods of time. This will lead to the same returns, on average, but the investor's net profits will be higher because fewer brokerage commissions will have to be paid.

A sensible strategy for a small investor, whose costs of managing a portfolio may be high relative to its size, is to buy into a mutual fund rather than to buy individual stocks. Because the efficient market hypothesis indicates that no mutual fund can consistently outperform the market, an investor should not buy into a fund that has high management fees or pays sales commissions to brokers but rather should purchase a no-load (commission-free) mutual fund that has low management fees. ◆

7.5 WHY THE EFFICIENT MARKET HYPOTHESIS DOES NOT IMPLY THAT FINANCIAL MARKETS ARE EFFICIENT

LO 7.5 Identify and explain the implications of the efficient market hypothesis for financial markets.

Many financial economists take the efficient market hypothesis one step further in their analysis of financial markets. Not only do they believe that expectations in financial markets are rational—that is, equal to optimal forecasts using all available information—but they also add the condition that prices in financial markets reflect the true fundamental (intrinsic) value of the securities. In other words, all prices are always correct and reflect **market fundamentals** (items that have a direct impact on future income streams of the securities), and so financial markets are efficient.

This stronger view of market efficiency has several important implications in the academic field of finance. First, it implies that in an efficient capital market, one investment is as good as any other because the securities' prices are correct. Second, it implies that a security's price reflects all available information about the intrinsic value of the security. Third, it implies that security prices can be used by managers of both financial and nonfinancial firms to assess their costs of capital (costs of financing their investments) accurately and hence that security prices can be used to help these managers make correct decisions about whether a specific investment is worth making. This stronger version of market efficiency is a basic tenet of much analysis in the finance field.

The efficient market hypothesis may be misnamed, however. It does not imply the stronger view of market efficiency but rather just that prices in markets like the stock market are unpredictable. Indeed, as the following application suggests, the existence of market crashes and **bubbles**, in which the prices of assets rise well above their fundamental values, casts serious doubt on the stronger view that financial markets are efficient. However, market crashes and bubbles do not necessarily provide strong evidence against the basic tenets of the efficient market hypothesis.

[7]The investor may also have to pay Revenue Canada capital gains taxes on any profits that are realized when a security is sold—an additional reason why continual buying and selling does not make sense.

APPLICATION

What Do Stock Market Crashes Tell Us About the Efficient Market Hypothesis and the Efficiency of Financial Markets?

On October 19, 1987, dubbed "Black Monday," the Dow Jones Industrial Average declined by more than 20%, the largest one-day decline in U.S. history. The collapse of the high-tech companies' share prices from their peaks in March 2000 caused the heavily tech-laden NASDAQ index to fall from about 5000 in March 2000 to about 1500 in 2001 and 2002, a decline of well over 60%. These stock market crashes caused many economists to question the validity of the efficient market hypothesis. These economists do not believe that a rational marketplace could have produced such a massive swing in share prices. To what degree should these stock market crashes make us doubt the validity of the efficient market hypothesis?

Nothing in efficient markets theory rules out large changes in stock prices. A large change in stock prices can result from new information that produces a dramatic decline in optimal forecasts of the future valuation of firms. However, economists are hard-pressed to find fundamental changes in the economy that would have caused the Black Monday and tech crashes. One lesson from these crashes is that factors other than market fundamentals probably have an effect on asset prices. Indeed, as we will explore in Chapters 8 and 11, there are good reasons to believe that impediments to the proper functioning of financial markets do exist. Hence these crashes have convinced many economists that the stronger version of market efficiency, which states that asset prices reflect the true fundamental (intrinsic) value of securities, is incorrect. These economists attribute a large role in determination of stock and other asset prices to market psychology and to the institutional structure of the marketplace. However, nothing in this view contradicts the basic reasoning behind rational expectations or the efficient market hypothesis—that market participants eliminate unexploited profit opportunities. Even though stock market prices may not always solely reflect market fundamentals, as long as market crashes are unpredictable, the basic premises of efficient markets theory hold.

However, other economists believe that market crashes and bubbles suggest that unexploited profit opportunities may exist and that the efficient market hypothesis might be fundamentally flawed. The controversy over the efficient market hypothesis continues. ◆

7.6 BEHAVIOURAL FINANCE

LO 7.6 Summarize the reasons why behavioural finance suggests that the efficient market hypothesis may not hold.

Doubts about the efficiency of financial markets, triggered by the stock market crash of 1987, led economists such as Nobel Prize winner Robert Shiller to develop a new field of study called **behavioural finance**. It applies concepts from other social sciences, such as anthropology, sociology, and particularly psychology, to explain the behaviour of securities prices.[8]

[8]Surveys of this field can be found in Hersh Shefrin, *Beyond Greed and Fear: Understanding of Behavioral Finance and the Psychology of Investing* (Boston: Harvard Business School Press, 2000); Andrei Shleifer, *Inefficient Markets* (Oxford, UK: Oxford University Press, 2000); and Robert J. Shiller, "From Efficient Market Theory to Behavioral Finance," Cowles Foundation Discussion Paper No. 1385 (October 2002).

As we have seen, the efficient market hypothesis assumes that unexploited profit opportunities are eliminated by "smart money" market participants. But can smart money dominate ordinary investors so that financial markets are efficient? Specifically, the efficient market hypothesis suggests that smart money participants will sell when a stock price goes up irrationally, with the result that the stock price falls back down to a level that is justified by fundamentals. For this to occur, smart money investors must be able to engage in **short sales**; that is, they must borrow stock from brokers and then sell it in the market, with the aim of earning a profit by buying the stock back again ("covering the short") after it has fallen in price. Work by psychologists, however, suggests that people are subject to loss aversion: They are more unhappy when they suffer losses than they are happy when they achieve gains. Short sales can result in losses far in excess of an investor's initial investment if the stock price climbs sharply higher than the price at which the short sale is made (and losses might be unlimited if the stock price climbs to astronomical heights).

Loss aversion can thus explain an important phenomenon: Very little short selling actually takes place. Short selling may also be constrained by rules restricting it, because it seems unsavoury for someone to make money from another person's misfortune. The existence of so little short selling can explain why stock prices are sometimes overvalued. That is, the lack of enough short selling means that smart money does not drive stock prices back down to their fundamental value.

Psychologists have also found that people tend to be overconfident in their own judgments. As a result, investors tend to believe that they are smarter than other investors. Because investors are willing to assume that the market typically doesn't get it right, they trade on their beliefs rather than on pure facts. This theory may explain why securities markets have such a large trading volume—something that the efficient market hypothesis does not predict.

Overconfidence and social contagion (fads) provide an explanation for stock market bubbles. When stock prices go up, investors attribute their profits to their intelligence and talk up the stock market. This word-of-mouth enthusiasm and glowing media reports then can produce an environment in which even more investors think stock prices will rise in the future. The result is a positive feedback loop in which prices continue to rise, producing a speculative bubble, which finally crashes when prices get too far out of line with fundamentals.[9]

The field of behavioural finance is a young one, but it holds out hope that we might be able to explain some features of securities markets' behaviour that are not well explained by the efficient market hypothesis.

[9]See Robert J. Shiller, *Irrational Exuberance* (New York: Broadway Books, 2001).

SUMMARY

1. Stocks are valued as the present value of future dividends. Unfortunately, we do not know very precisely what these dividends will be. This uncertainty introduces a great deal of error into the valuation process. The Gordon growth model is a simplified method of computing stock value that depends on the assumption that the dividends are growing at a constant rate forever. Given our uncertainty regarding future dividends, this assumption is often the best we can do.
2. The interaction among traders in the market is what actually sets prices on a day-to-day basis. The trader who values the security the most (either because of less uncertainty about cash flows or because of greater estimated cash flows) will be willing to pay the most. As new information is released, investors will revise their estimates of the true value of the security and will either buy or sell it, depending on how the market price compares with their estimated valuation. Because small

changes in estimated growth rates or required returns result in large changes in price, it is not surprising that the markets are often volatile.

3. The efficient market hypothesis states that current security prices will fully reflect all available information, because in an efficient market, all unexploited profit opportunities are eliminated. The elimination of unexploited profit opportunities necessary for a financial market to be efficient does not require that all market participants be well informed. The efficient markets hypothesis implies that stock prices generally follow a random walk.
4. The efficient market hypothesis indicates that hot tips and investment advisers' published recommendations cannot help an investor outperform the market. The best prescription for investors is to pursue a buy-and-hold strategy—purchase stocks and hold them for long periods of time. Empirical evidence generally supports these implications of the efficient market hypothesis in the stock market.
5. The existence of market crashes and bubbles has convinced many economists that the stronger version of market efficiency, which states that asset prices reflect the true fundamental (intrinsic) value of securities, is not correct. There is, however, less evidence that these crashes prove that the efficient market hypothesis is wrong. Even if the stock market were driven by factors other than fundamentals, these crashes do not clearly demonstrate that many basic tenets of the efficient market hypothesis are no longer valid, as long as the crashes could not have been predicted.
6. The new field of behavioural finance applies concepts from other social sciences, such as anthropology, sociology, and psychology, to our understanding of the behaviour of securities prices. Loss aversion, overconfidence, and social contagion can explain why trading volume is so high, why stock prices become overvalued, and why speculative bubbles occur.

MyLab Economics Make the grade with MyLab Economics: Select end-of-chapter questions can be found on MyLab Economics. You can practise them as often as you want, and most feature step-by-step guided instructions to help you find the right answer.

QUESTIONS

1. What basic principle of finance can be applied to the valuation of any investment asset?
2. What are the two main sources of cash flows for a stockholder? How reliably can these cash flows be estimated? Compare the problem of estimating stock cash flows to the problem of estimating bond cash flows. Which security would you predict to be more volatile?
3. Some economists think that central banks should try to prick bubbles in the stock market before they get out of hand and cause later damage when they burst. How can monetary policy be used to prick a market bubble? Explain using the Gordon growth model.
4. If monetary policy becomes more transparent about the future course of interest rates, how will stock prices be affected, if at all?
5. Suppose that you are asked to forecast future stock prices of ABC Corporation, so you proceed to collect all available information. The day you announce your forecast, competitors of ABC Corporation announce a brand-new plan to merge and reshape the structure of the industry. Would your forecast still be considered optimal?
6. "Anytime it is snowing when Joe Commuter gets up in the morning, he misjudges how long it will take him to drive to work. When it is not snowing, his expectations of the driving time are perfectly accurate. Considering that it snows only once every 10 years where Joe lives, Joe's expectations are almost always perfectly accurate." Are Joe's expectations rational? Why or why not?
7. Suppose that you decide to play a game. You buy stock by throwing a die a few times, using that method to select which stock to buy. After 10 months you calculate the

return on your investment and the return earned by someone who followed "expert" advice during the same period. If both returns are similar, would this constitute evidence in favour of or against the efficient market hypothesis?

8. "If stock prices did not follow a random walk, there would be unexploited profit opportunities in the market." Is this statement true, false, or uncertain? Explain your answer.
9. Suppose that increases in the money supply lead to a rise in stock prices. Does this mean that when you see that the money supply has sharply increased in the past week, you should go out and buy stocks? Why or why not?
10. If the public expects a corporation to lose $5 per share this quarter and it actually loses $4, which is still the largest loss in the history of the company, what does the efficient market hypothesis predict will happen to the price of the stock when the $4 loss is announced?
11. If you read in the *Globe and Mail* that the "smart money" on Bay Street expects stock prices to fall, should you follow that lead and sell all your stocks?
12. If your broker has been right in their five previous buy and sell recommendations, should you continue listening to their advice?
13. Can a person with rational expectations expect the price of a share of Google to rise by 10% in the next month?
14. Suppose that in every last week of November stock prices go up by an average of 3%. Would this constitute evidence in favour of or against the efficient market hypothesis?
15. "An efficient market is one in which no one ever profits from having better information than the rest of the market participants." Is this statement true, false, or uncertain? Explain your answer.
16. If higher money growth is associated with higher future inflation, and if announced money growth turns out to be extremely high but is still less than the market expected, what do you think will happen to long-term bond prices?
17. "Foreign exchange rates, like stock prices, should follow a random walk." Is this statement true, false, or uncertain? Explain your answer.
18. Assume that the efficient market hypothesis holds. Marcos has been recently hired by a brokerage firm and claims that he now has access to the best market information. However, he is the "new guy," and no one at the firm tells him much about the business. Would you expect Marcos's clients to be better or worse off than the rest of the firm's clients?
19. Suppose that you work as a forecaster of future monthly inflation rates and that your last six forecasts have been off by minus 1%. Is it likely that your expectations are optimal?
20. In the late 1990s, as information technology advanced rapidly and the Internet became widely developed, U.S. stock markets soared, peaking in early 2001. Later that year, these markets began to unwind and then crashed, with many commentators identifying the previous few years as a "stock market bubble." How might it be possible for this episode to be a bubble but still adhere to the efficient market hypothesis?
21. Why might the efficient market hypothesis be less likely to hold when fundamentals suggest stocks should be at a lower level?

APPLIED PROBLEMS

22. Compute the price of a share of stock that pays a $1 per year dividend and that you expect to be able to sell in one year for $20, assuming you require a 15% return.
23. After careful analysis, you have determined that a firm's dividends should grow at 7%, on average, in the foreseeable future. The firm's last dividend was $3. Compute the current price of this stock, assuming the required return is 18%.
24. The current price of a stock is $65.88. If dividends are expected to be $1 per share for the next five years, and the required return is 10%, what should the price of the stock be in five years when you plan to sell it? If the dividend and required return remain the same, and the stock price is expected to increase by $1 five years from now, does the current stock price also increase by $1? Why or why not?

25. A company has just announced a 3-for-1 stock split, effective immediately. Prior to the split, the company had a market value of $5 billion with 100 million shares outstanding. Assuming the split conveys no new information about the company, what are the value of the company, the number of shares outstanding, and the price per share after the split? If the actual market price immediately following the split is $17.00 per share, what does this tell us about market efficiency?

DATA ANALYSIS PROBLEMS

1. Go to the CANSIM database and download monthly data from January 1956 to December 2020 on the S&P TSX Composite Index (series V122620).
   a. Calculate the mean and standard deviation as well as the maximum and minimum values for this series.
   b. Which were the worst and best months in terms of rates of return?

2. **Real-Time Data Analysis** Go to the St. Louis Federal Reserve FRED database and find data on the Dow Jones Industrial Average (DJIA). Assume the DJIA is a stock that pays no dividends. Apply the one-period valuation model, using the data from one year prior up to the most current date available, to determine the required return on equity investment. In other words, assume the most recent stock price of DJIA is known one year prior. What rate of return would be required to "buy" a share of DJIA? Suppose that a 100 points dividend is paid out instead. How does this change the required rate of return?

PART 3

Financial Institutions

Crisis and Response: Bailout Packages in the Trillions of Dollars

In response to the global financial crisis that began in August 2007, governments and central banks around the world departed significantly from their traditional policy tools and introduced new facilities, pouring trillions of dollars into financial institutions. In the United States, for example, the House of Representatives passed the Emergency Economic Stabilization Act on October 3, 2008. This stunning $700 billion bailout package sought to promote recovery from the financial crisis by authorizing the U.S. Treasury to purchase troubled mortgage assets from struggling financial institutions or to inject capital into banking institutions. To calm fears further, the act raised the federal deposit insurance limit from $100 000 to $250 000.

The initial bill was voted down on September 29 when constituents flooded their representatives with complaints about bailing out the greedy Wall Street executives behind the crisis. The debate in the United States pitted Wall Street against Main Street: Many people who sided with struggling homeowners saw the proposed federal bailout of financial institutions as government hypocrisy. How could injecting capital into the financial system help those fearful of losing their job or, worse yet, those who found themselves suddenly without work?

The central role of financial institutions in the working of the economy—the focus of Part III—was being overlooked. Banks and other financial institutions make financial markets work by moving funds from people who save to people who have productive investment opportunities. That bank branch on Main Street was not going to be able to lend freely to a small business owner or recent university graduate looking to fund a new car purchase until capital once again flowed.

The global financial crisis highlights how the financial systems change over time, be it from financial innovations or through hard lessons learned from such crises. Chapter 8 analyzes financial structures in Canada and in the rest of the world. In Chapter 9, we extend the economic analysis developed in Chapter 8 to understand the motivations for bank regulation, and we examine the pitfalls in the regulatory process. Chapter 10 examines the development of the Canadian banking system over time and the growing internationalization of banking. In Chapter 11, we develop a framework for understanding the dynamics of financial crises, which is then used to explain what happened during the global financial crisis of 2007–2009 and whether the coronavirus pandemic could have led to another financial crisis.

8 An Economic Analysis of Financial Structure

Learning Objectives

8.1 Identify eight basic facts about the global financial system.

8.2 Summarize how transaction costs affect financial intermediaries.

8.3 Describe why asymmetric information leads to adverse selection and moral hazard.

8.4 Recognize adverse selection and summarize the ways in which it can be reduced.

8.5 Recognize the principal–agent problem arising from moral hazard in equity contracts and summarize the methods for reducing it.

8.6 Summarize the methods used to reduce moral hazard in debt contracts.

Preview

A healthy and vibrant economy requires a financial system that moves funds from people who save to people who have productive investment opportunities. But how does the financial system make sure that your hard-earned savings get channelled to Paula the Productive Investor rather than to Benny the Bum?

This chapter answers that question by providing an economic analysis of how our financial structure is designed to promote economic efficiency. The analysis focuses on a few simple but powerful economic concepts that enable us to explain features of our financial system, such as why financial contracts are written as they are and why financial intermediaries are more important than securities markets for getting funds to borrowers. The analysis also demonstrates the important link between the financial system and the performance of the aggregate economy, which is the subject of Part 6 of this book.

8.1 BASIC FACTS ABOUT FINANCIAL STRUCTURE THROUGHOUT THE WORLD

LO 8.1 Identify eight basic facts about the global financial system.

The financial system is complex in structure and function throughout the world. It includes many different types of institutions: banks, insurance companies, mutual funds, stock and bond markets, and so on—all of which are regulated by government. The financial system channels trillions of dollars per year from savers to people with productive investment opportunities. If we take a close look at financial structure all over the world, we find eight basic facts, some of which are quite surprising, that we must explain if we are to understand how the financial system works.

The bar chart in Figure 8-1 shows how Canadian businesses financed their activities using external funds (those obtained from outside the business itself) in the period 1970–2000 and compares Canadian data with data for the United States, Germany, and Japan. The conclusions drawn from this period still hold true today. The *Bank Loans* category is made up primarily of loans from depository institutions; *Nonbank Loans* are primarily loans by other financial intermediaries; the *Bonds* category includes marketable debt securities, such as corporate bonds and commercial paper; and *Stock* consists of new issues of equity (stock market shares).

Mini-lecture

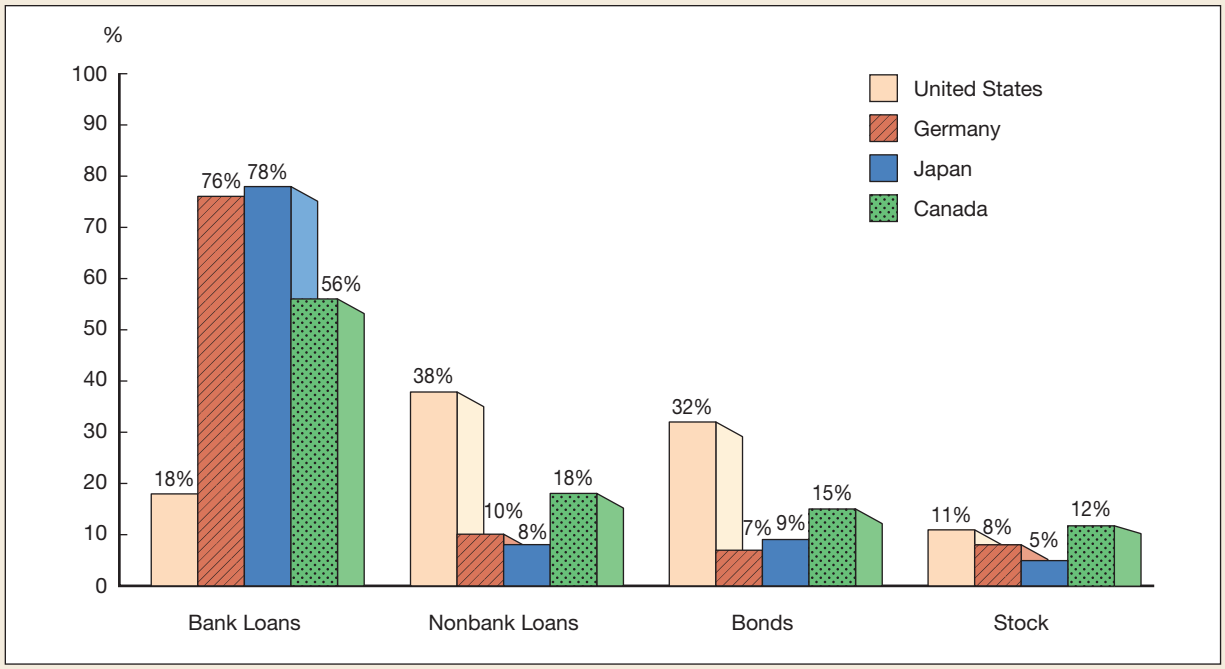


FIGURE 8-1 Sources of External Funds for Nonfinancial Businesses: A Comparison of Canada with Germany, Japan, and the United States

The *Bank Loans* category is made up primarily of loans from depository institutions; the *Nonbank Loans* are primarily loans by other financial intermediaries; the *Bonds* category includes marketable debt securities, such as corporate bonds and commercial paper; and *Stock* consists of new issues of new equity (stock market shares).
Source: Andreas Hackethal and Reinhard H. Schmidt, "Financing Patterns: Measurement Concepts and Empirical Results," Johann Wolfgang Goethe-Universitat Working Paper No. 125, January 2004. The data are for 1970–2000 and are gross flows as percentages of the total, not including trade and other credit data, which are not available.

Now let's explore the eight facts.

1. ***Stocks are not the most important source of external financing for businesses.*** Because so much attention in the media is focused on the stock market, many people have the impression that stocks are the most important source of financing for Canadian corporations. However, as we can see from the bar chart in Figure 8-1, the stock market accounted for only a small fraction of the external financing of Canadian businesses in the 1970–2000 period: 12%.[1] Similarly low figures apply for the other countries presented in Figure 8-1. Why is the stock market less important than other sources of financing in the United States and other countries?

[1]The 12% figure for the percentage of external financing provided by stocks is based on the flows of external funds to corporations. However, this flow figure is somewhat misleading because when a share of stock is issued, it raises funds permanently, whereas when a bond is issued, it raises funds only temporarily until it is paid back at maturity. To see this, suppose a firm raises $1000 by selling a share of stock and another $1000 by selling a $1000 one-year bond. In the case of the stock issue, the firm can hold on to the $1000 it raised by selling the stock. However, to hold on to the $1000 it raised through debt, the firm has to issue a new $1000 bond every year. If we look at the flow of funds to corporations over a 30-year period, as shown in Figure 8-1, the firm will have raised $1000 through a stock issue only once in the 30-year period, while it will have raised $1000 through debt 30 times, once in each of the 30 years. Thus it looks as though debt is 30 times more important than stocks in raising funds, even though our example indicates that the two methods are actually equally important for the firm.

2. ***Issuing marketable debt and equity securities is not the primary way in which businesses finance their operations.*** Figure 8-1 shows that bonds are a more important source of financing in Canada than stocks are (15% versus 12%). However, stocks and bonds combined (27%), which make up the total share of marketable securities, still supply about a bit over 25% of the external funds needed by Canadian corporations to finance their activities. The fact that issuing marketable securities is not the most important source of financing is true elsewhere in the world as well. Indeed, as we see in Figure 8-1, other countries (except the United States) have a smaller share of external financing supplied by marketable securities than does Canada. Why don't businesses use marketable securities more extensively to finance their activities?
3. ***Indirect finance, which involves the activities of financial intermediaries, is many times more important than direct finance, in which businesses raise funds directly from lenders in financial markets.*** Direct finance involves the sale to households of marketable securities, such as stocks and bonds. The 27% share of stocks and bonds used as a source of external financing for Canadian businesses actually greatly overstates the importance of direct finance in our financial system. In general, only a small fraction of newly issued corporate bonds, commercial paper, and new issues of stocks have been sold directly to Canadian households. The rest of these securities have been bought primarily by financial intermediaries such as insurance companies, pension funds, and mutual funds. Because in most countries marketable securities are an even less important source of finance than in Canada, direct finance is also far less important than indirect finance in the rest of the world. Why are financial intermediaries and indirect finance so important in financial markets? In recent years, however, indirect finance has been declining in importance. Why is this happening?
4. ***Financial intermediaries, particularly banks, are the most important source of external funds used to finance businesses.*** As we can see in Figure 8-1, the primary source of external funds for businesses throughout the world is loans made by banks and other nonbank financial intermediaries, such as insurance companies, pension funds, and finance companies (56% in Canada, but more than 70% in Germany and Japan). In other industrialized countries, bank loans are the largest category of sources of external finance (more than 70% in Germany and Japan). Thus the data suggest that banks in these countries have the most important role in financing business activities. In developing countries, banks play an even more important role in the financial system than they do in industrialized countries. What makes banks so important to the workings of the financial system? Although banks remain important, their contributions to external funds for businesses have been declining in recent years. What is driving this decline?
5. ***The financial system is among the most heavily regulated sectors of the economy.*** The financial system is heavily regulated in Canada and in all other developed countries. Governments regulate financial markets primarily to promote the provision of information and to ensure the soundness (stability) of the financial system. Why are financial markets so extensively regulated throughout the world?
6. ***Only large, well-established corporations have easy access to securities markets to finance their activities.*** Individuals and smaller businesses that are not well established are less likely to raise funds by issuing marketable securities. Instead, they most often obtain their financing from banks. Why do only large, well-known corporations find it easier to raise funds in securities markets?

7. ***Collateral is a prevalent feature of debt contracts for both households and businesses.*** **Collateral** is property that is pledged to a lender to guarantee payment in the event that the borrower is unable to make debt payments. Collateralized debt (also known as **secured debt** to contrast it with **unsecured debt**, such as credit card debt, which is not collateralized) is the predominant form of household debt and is widely used in business borrowing as well. The majority of household debt in Canada consists of collateralized loans: Your automobile is collateral for your auto loan, and your house is collateral for your mortgage. Commercial and farm mortgages, for which property is pledged as collateral, make up one-quarter of borrowing by nonfinancial businesses; corporate bonds and other bank loans also often involve pledges of collateral. Why is collateral such an important feature of debt contracts?
8. ***Debt contracts typically are extremely complicated legal documents that place substantial restrictions on the behaviour of the borrower.*** Many students think of a debt contract as a simple IOU that can be written on a single piece of paper. The reality of debt contracts is far different, however. In all countries, bond or loan contracts typically are long legal documents with provisions (called **restrictive covenants**) that restrict and specify certain activities that the borrower can engage in. Restrictive covenants are not just a feature of debt contracts for businesses; for example, personal automobile loan and home mortgage contracts include covenants that require the borrower to maintain sufficient insurance on the automobile or house purchased with the loan. Why are debt contracts so complex and restrictive?

As you may recall from Chapter 2, an important feature of financial markets is their substantial transaction and information costs. An economic analysis of how these costs affect financial markets will help us understand the eight facts, which will in turn give us a much deeper understanding of how our financial system works. In the next section, we examine the impact of transaction costs on the structure of our financial system. Then we turn to the effects of information costs on financial structure.

8.2 TRANSACTION COSTS

LO 8.2 Summarize how transaction costs affect financial intermediaries.

Transaction costs are a major problem in financial markets. An example will make this clear.

How Transaction Costs Influence Financial Structure

Say you have $5000 that you would like to invest, and you are thinking about investing in the stock market. Because you have only $5000, you can buy only a small number of shares. Even if you use online trading, your purchase is so small that the brokerage commission for buying the stock you pick will be a large percentage of the purchase price of the shares. If, instead, you decide to buy a bond, the problem becomes even worse; the smallest denomination offered on some bonds that you might want to buy is as large as $10 000, and you do not have that much money to invest. You are disappointed and realize that you will not be able to use financial markets to earn a return on your hard-earned savings. You can take some consolation, however, in the fact that you are not alone in being stymied by high transaction costs. This is a fact of life for many of us: About one-half of Canadian households own any securities.

You also face another problem related to transaction costs. Because you have only a small amount of funds available, you can make only a restricted number of

investments, because a large number of small transactions would result in very high transaction costs. That is, you have to put all your eggs in one basket, and your inability to diversify will subject you to a lot of risk.

How Financial Intermediaries Reduce Transaction Costs

This example of the problems posed by transaction costs, along with the example outlined in Chapter 2, in which legal costs kept you from making a loan to Carl the Carpenter, illustrates that small savers like you are frozen out of financial markets and are unable to benefit from them. Fortunately, financial intermediaries, an important part of the financial structure, have evolved to reduce transaction costs and allow small savers and borrowers to benefit from the existence of financial markets.

Economies of Scale One solution to the problem of high transaction costs is to bundle the funds of many investors together so that they can take advantage of *economies of scale*, the reduction in transaction costs per dollar of investment as the size (scale) of transactions increases. Bundling investors' funds together reduces transaction costs for each individual investor. Economies of scale exist because the total cost of carrying out a transaction in financial markets increases only a little as the size of the transaction grows. For example, the cost of arranging a purchase of 10 000 shares of stock is not much greater than the cost of arranging a purchase of 50 shares of stock.

The presence of economies of scale in financial markets helps explain the development of financial intermediaries and why financial intermediaries have become such an important part of our financial structure. The clearest example of a financial intermediary that arose because of economies of scale is a mutual fund. A *mutual fund* is a financial intermediary that sells shares to individuals and then invests the proceeds in bonds or stocks. Because it buys large blocks of stocks or bonds, a mutual fund can take advantage of lower transaction costs. These cost savings are then passed on to individual investors after the mutual fund has taken its cut in the form of management fees for administering their accounts. An additional benefit for individual investors is that a mutual fund is large enough to purchase a widely diversified portfolio of securities. The increased diversification for individual investors reduces their risk, making them better off.

Economies of scale are also important in lowering the costs of resources that financial institutions need to accomplish their tasks, such as computer technology. Once a large mutual fund has invested a lot of money in setting up a telecommunications system, for example, the system can be used for a huge number of transactions at a low cost per transaction.

Expertise Financial intermediaries are also better able to develop expertise that can be used to lower transaction costs. Their expertise in computer technology, for example, enables them to offer their customers convenient services such as cheque-writing privileges on their accounts and toll-free numbers that customers can call for information on how well their investments are doing.

Low transaction costs enable financial intermediaries to provide their customers with *liquidity services*, which are services that make it easier for customers to conduct transactions. Money market mutual funds, for example, not only pay shareholders relatively high interest rates but also allow them to write cheques for convenient bill paying.

8.3 ASYMMETRIC INFORMATION: ADVERSE SELECTION AND MORAL HAZARD

LO 8.3 Describe why asymmetric information leads to adverse selection and moral hazard.

The presence of transaction costs in financial markets partly explains why financial intermediaries and indirect finance play such an important role in financial markets (fact 3). To understand financial structure more fully, however, we turn to the role of information in financial markets.

Asymmetric information—a situation that arises when one party's insufficient knowledge about the other party involved in a transaction makes it impossible for the first party to make accurate decisions when conducting the transaction—is an important aspect of financial markets. For example, managers of a corporation know whether they are honest and usually have better information about how well their business is doing than stockholders do. The presence of asymmetric information leads to adverse selection and moral hazard problems, which were introduced in Chapter 2.

Adverse selection is an asymmetric information problem that occurs *before* a transaction occurs. Adverse selection occurs when one party to a transaction has information about a hidden characteristic and takes economic advantage of this information by engaging in a transaction with less informed parties. Specifically, potential bad credit risks are the ones who most actively seek out loans. Thus the parties who are most likely to produce an undesirable outcome are also the ones most likely to want to engage in the transaction. For example, big risk takers or outright crooks are often the most eager to take out a loan because they know they are unlikely to pay it back. Because adverse selection increases the chances that a loan might be made to a bad credit risk, lenders might decide not to make any loans, even though good credit risks can be found in the marketplace.

Moral hazard arises *after* the transaction occurs. Moral hazard occurs when an informed party takes a hidden (unobserved) action that harms the less informed party. A lender runs the risk that the borrower will engage in activities that are undesirable from the lender's point of view, because such activities make it less likely that the loan will be paid back. For example, once borrowers have obtained a loan, they may take on big risks (which have possible high returns but also run a greater risk of default) because they are playing with someone else's money. Because moral hazard lowers the probability that the loan will be repaid, lenders may decide that they would rather not make a loan.

The analysis of how asymmetric information problems affect economic behaviour is called **agency theory**. We will apply this theory here to explain why financial structure takes the form it does, thereby explaining the facts outlined at the beginning of the chapter.

8.4 THE LEMONS PROBLEM: HOW ADVERSE SELECTION INFLUENCES FINANCIAL STRUCTURE

LO 8.4 Recognize adverse selection and summarize the ways in which it can be reduced.

A particular aspect of the way the adverse selection problem interferes with the efficient functioning of a market was outlined in a famous article by Nobel Prize winner George Akerlof. It is called the "lemons problem" because it resembles the problem created by

"lemons," that is, bad cars, in the used-car market.[2] Potential buyers of used cars are frequently unable to assess the quality of a car; that is, they can't tell whether a particular used car is one that will run well or a lemon that will continually give them grief. The price that a buyer pays must therefore reflect the *average* quality of the cars in the market, somewhere between the low value of a lemon and the high value of a good car.

The owner of a used car, by contrast, is more likely to know whether the car is a peach or a lemon. If the car is a lemon, the owner is more than happy to sell it at the price the buyer is willing to pay, which, being somewhere between the value of a lemon and that of a good car, is greater than the lemon's value. However, if the car is a peach, that is, a good car, the owner knows that the car is undervalued at the price the buyer is willing to pay, and so the owner may not want to sell it. As a result of this adverse selection problem, fewer good used cars will come to the market. Because the average quality of a used car available in the market will be low, and because very few people want to buy a lemon, there will be few sales. The used-car market will function poorly, if at all.

Lemons in the Stock and Bond Markets

A similar lemons problem arises in securities markets—that is, the debt (bond) and equity (stock) markets. Suppose that our friend Irving the Investor, a potential buyer of securities such as common stock, can't distinguish between good firms with high expected profits and low risk and bad firms with low expected profits and high risk. In this situation, Irving will be willing to pay only a price that reflects the *average* quality of firms issuing securities—a price that lies between the value of securities from bad firms and the value of those from good firms. If the owners or managers of a good firm have better information than Irving and *know* that they have a good firm, then they know that their securities are undervalued and will not want to sell them to Irving at the price he is willing to pay. The only firms willing to sell Irving securities will be bad firms (because his price is higher than the securities are worth). Our friend Irving is not stupid; he does not want to hold securities in bad firms, and hence he will decide not to purchase securities in the market. In an outcome similar to that in the used-car market, this securities market will not work very well because few firms will sell securities in it to raise capital.

The analysis is similar if Irving considers purchasing a corporate debt instrument in the bond market rather than an equity share. Irving will buy a bond only if its interest rate is high enough to compensate him for the average default risk of the good and bad firms trying to sell the debt. The knowledgeable owners of a good firm realize that they will be paying a higher interest rate than they should, so they are unlikely to want to borrow in this market. Only the bad firms will be willing to borrow, and because investors like Irving are not eager to buy bonds issued by bad firms, they will probably not buy any bonds at all. Few bonds are likely to sell in this market, so it will not be a good source of financing.

The analysis we have just conducted explains fact 2—why marketable securities are not the primary source of financing for businesses in any country in the world. It also partly explains fact 1—why stocks are not the most important source of financing for Canadian businesses. The presence of the lemons problem keeps securities markets such as the stock and bond markets from being effective in channelling funds from savers to borrowers.

[2]George Akerlof, "The Market for 'Lemons': Quality, Uncertainty and the Market Mechanism," *Quarterly Journal of Economics* 84 (1970): 488–500. Two important papers that have applied the lemons problem analysis to financial markets are Stewart Myers and N. S. Majluf, "Corporate Financing and Investment Decisions: When Firms Have Information That Investors Do Not Have," *Journal of Financial Economics* 13 (1984): 187–221; and Bruce Greenwald, Joseph E. Stiglitz, and Andrew Weiss, "Information Imperfections in the Capital Market and Macroeconomic Fluctuations," *American Economic Review* 74 (1984): 194–199.

Tools to Help Solve Adverse Selection Problems

In the absence of asymmetric information, the lemons problem goes away. If buyers know as much about the quality of used cars as sellers, so that all involved can tell a good car from a bad one, buyers will be willing to pay full value for good used cars. Because the owners of good used cars can now get a fair price, they will be willing to sell them in the market. The market will have many transactions and will perform its intended job of channelling good cars to people who want them.

Similarly, if purchasers of securities can distinguish good firms from bad, they will pay the full value of securities issued by good firms, and good firms will sell their securities in the market. The securities market will then be able to move funds to the good firms that have the most productive investment opportunities.

Private Production and Sale of Information

The solution to the adverse selection problem in financial markets is to reduce asymmetric information by furnishing the people supplying funds with more details about the individuals or firms seeking to finance their investment activities. One way for saver-lenders to get this information is through private companies that collect and produce information distinguishing good firms from bad firms and then sell it to the saver-lenders. In Canada, companies such as Standard & Poor's and the Dominion Bond Rating Service gather information on firms' balance sheet positions and investment activities, publish these data, and sell them to subscribers (individuals, libraries, and financial intermediaries involved in purchasing securities).

The system of private production and sale of information does not completely solve the adverse selection problem in securities markets, however, because of the **free-rider problem**. The free-rider problem occurs when people who do not pay for information take advantage of the information that other people have paid for. The free-rider problem suggests that the private sale of information is only a partial solution to the lemons problem. To see why, suppose you have just purchased information that tells you which firms are good and which are bad. You believe that this purchase is worthwhile because you can make up the cost of acquiring this information, and then some, by purchasing the securities of good firms that are undervalued. However, when our savvy (free-riding) investor Irving sees you buying certain securities, he buys right along with you, even though he has not paid for any information. If many other investors act as Irving does, the increased demand for the undervalued good securities causes their low price to be bid up immediately to reflect the securities' true value. Because of all these free riders, you can no longer buy the securities for less than their true value. Now, because you will not gain any profit from purchasing the information, you realize that you never should have paid for the information in the first place. If other investors come to the same realization, private firms and individuals may not be able to sell enough of this information to make it worth their while to gather and produce it. The weakened ability of private firms to profit from selling information will mean that less information is produced in the marketplace, so adverse selection (the lemons problem) will still interfere with the efficient functioning of securities markets.

Government Regulation to Increase Information

The free-rider problem prevents the private market from producing enough information to eliminate all the asymmetric information that leads to adverse selection. Could financial markets benefit from government intervention? The government could, for instance, produce information to help investors distinguish good from bad firms and provide it to the public free of charge. This solution, however, would involve the government releasing negative information about firms, a practice that might be politically difficult. A second possibility (and one followed by Canada and most governments throughout the world) is for the government

FYI The Enron Implosion

Until 2001, Enron Corporation, a firm that specialized in trading in the energy market, appeared to be spectacularly successful. It controlled a quarter of the energy-trading market and was valued as high as $77 billion in August 2000 (just a little over a year before its collapse), making it the seventh-largest corporation in the United States at that time. However, toward the end of 2001, Enron came crashing down. In October 2001, Enron announced a third-quarter loss of $618 million and disclosed accounting "mistakes." The SEC then engaged in a formal investigation of Enron's financial dealings with partnerships led by its former finance chief. It became clear that Enron was engaged in a complex set of transactions by which it was keeping substantial amounts of debt and financial contracts off its balance sheet. These transactions enabled Enron to hide its financial difficulties. Despite securing as much as $1.5 billion of new financing from J. P. Morgan Chase and Citigroup, the company was forced to declare bankruptcy in December 2001, up to that point the largest bankruptcy declaration in U.S. history.

The Enron collapse illustrates that government regulation can lessen asymmetric information problems but cannot eliminate them. Managers have tremendous incentives to hide their companies' problems, making it hard for investors to know the true value of firms.

The Enron bankruptcy not only increased concerns in financial markets about the quality of accounting information supplied by corporations but also led to hardship for many of the firm's former employees, who found that their pensions had become worthless. Outrage against the duplicity of executives at Enron was high, and several of them were convicted and sent to jail.

to regulate securities markets in a way that encourages firms to reveal honest information about themselves so that investors can determine how good or bad the firms are. In Canada, government regulations require firms selling their securities to undergo independent audits, in which accounting firms certify that the firm is adhering to standard accounting principles and disclosing accurate information about sales, assets, and earnings. Similar regulations are found in other countries. However, disclosure requirements do not always work well, as the collapse of Enron and accounting scandals at other corporations, such as WorldCom and Parmalat (an Italian company), suggest (see the FYI box "The Enron Implosion").

The asymmetric information problem of adverse selection in financial markets helps explain why financial markets are among the most heavily regulated sectors of the economy (fact 5). Government regulation aimed at increasing the information available to investors is necessary to reduce the adverse selection problem, which interferes with the efficient functioning of securities (stock and bond) markets.

Although government regulation lessens the adverse selection problem, it does not eliminate it entirely. Even when firms provide information to the public about their sales, assets, or earnings, they still have more information than investors: A lot more is involved in knowing the quality of a firm than statistics alone can provide. Furthermore, bad firms have an incentive to make themselves look like good firms because this enables them to fetch a higher price for their securities. Bad firms will slant the information they are required to transmit to the public, thus making it harder for investors to sort out the good firms from the bad.

Financial Intermediation So far we have seen that private production of information and government regulation to encourage provision of information lessen, but do not eliminate, the adverse selection problem in financial markets. How, then, can

the financial structure help promote the flow of funds to people with productive investment opportunities when asymmetric information exists? A clue is provided by the structure of the used-car market.

An important feature of the used-car market is that most used cars are not sold directly by one individual to another. An individual who considers buying a used car might pay for privately produced information by subscribing to a magazine like *Consumer Reports* to find out if a particular make of car has a good repair record. Nevertheless, reading *Consumer Reports* does not solve the adverse selection problem, because even if a particular make of car has a good reputation, the specific car someone is trying to sell could be a lemon. The prospective buyer might also bring the used car to a mechanic for a once-over. But what if the prospective buyer doesn't know a mechanic who can be trusted or the mechanic charges a high fee to evaluate the car?

Because these roadblocks make it hard for individuals to acquire enough information about used cars, most used cars are not sold directly by one individual to another. Instead, they are sold by an intermediary, a used-car dealer who purchases used cars from individuals and resells them to other individuals. Used-car dealers produce information in the market by becoming experts in determining whether a car is a peach or a lemon. Once a dealer knows that a car is good, the dealer can sell it with some form of a guarantee: either an explicit guarantee such as a warranty or an implicit guarantee in which the dealer stands by its reputation for honesty. People are more likely to purchase a used car because of a dealer guarantee, and the dealer is able to sell the used car at a higher price than they paid for it. Thus the dealer profits from the production of information about automobile quality. If dealers purchase and then resell cars on which they have produced information, they avoid the problem of other people free-riding on the information they produced.

Just as used-car dealers help solve adverse selection problems in the automobile market, financial intermediaries play a similar role in financial markets. A financial intermediary, such as a bank, becomes an expert in producing information about firms so that it can sort out good credit risks from bad ones. It then can acquire funds from depositors and lend them to the good firms. Because the bank is able to lend mostly to good firms, it is able to earn a higher return on its loans than the interest it has to pay to its depositors. The resulting profit that the bank earns gives it the incentive to engage in this information production activity.

An important element of the bank's ability to profit from the information it produces is that it avoids the free-rider problem by primarily making private loans rather than by purchasing securities that are traded in the open market. Because a private loan is not traded, other investors cannot watch what the bank is doing and bid down the loan's interest rate to the point that the bank receives no compensation for the information it has produced. The bank's role as an intermediary that holds mostly nontraded loans is the key to its success in reducing asymmetric information in financial markets.

Our analysis of adverse selection indicates that financial intermediaries in general—and banks in particular, because they hold a large fraction of nontraded loans—should play a greater role in moving funds to corporations than securities markets do. Our analysis thus explains facts 3 and 4: why indirect finance is so much more important than direct finance and why banks are the most important source of external funds for financing businesses.

Our analysis also explains the greater importance of banks, as opposed to securities markets, in the financial systems of developing countries. As we have seen, better information about the quality of firms lessens asymmetric information problems, making it easier for firms to issue securities. Information about private firms is harder to collect in

developing countries than in industrialized countries; therefore, the smaller role played by securities markets leads to a greater role for financial intermediaries such as banks. As a corollary to our analysis, as information about firms becomes easier to acquire, the role of banks should decline. A major development in the past 30 years in Canada has been huge improvements in information technology. Thus our analysis suggests that the lending role of financial institutions such as banks in Canada should have declined, and this is exactly what has occurred (see Chapter 10).

Our analysis of adverse selection also explains fact 6, which questions why large firms are more likely to obtain funds from securities markets, a direct route, rather than from banks and financial intermediaries, an indirect route. The better known a corporation is, the more information about its activities is available in the marketplace. Thus it is easier for investors to evaluate the quality of the corporation and determine whether it is a good firm or a bad one. Because investors have fewer worries about adverse selection when dealing with well-known corporations, they are more willing to invest directly in their securities. Thus, in accordance with adverse selection, a pecking order for firms that can issue securities should exist. Hence we have an explanation for fact 6: The larger and more established a corporation is, the more likely it will be to issue securities to raise funds.

Collateral and Net Worth Adverse selection interferes with the functioning of financial markets only if a lender suffers a loss when a borrower is unable to make loan payments and thereby defaults on the loan. *Collateral*, property promised to the lender if the borrower defaults, reduces the consequences of adverse selection because it reduces the lender's losses in the event of a default. If a borrower defaults on a loan, the lender can sell the collateral and use the proceeds to make up for the losses on the loan. For example, if you fail to make your mortgage payments, the lender can take the title to your house, auction it off, and use the receipts to pay off the loan. Lenders are thus more willing to make loans secured by collateral, and borrowers are willing to supply collateral because the reduced risk for the lender makes it more likely that the loan will be made, perhaps even at a better loan rate. The presence of adverse selection in credit markets thus explains why collateral is an important feature of debt contracts (fact 7).

Net worth (also called **equity capital**), the difference between a firm's assets (what it owns or is owed) and its liabilities (what it owes), can perform a similar role to that of collateral. If a firm has a high net worth, then even if it engages in investments that lead to negative profits and so defaults on its debt payments, the lender can take title to the firm's net worth, sell it off, and use the proceeds to recoup some of the losses from the loan. In addition, the more net worth a firm has in the first place, the less likely it is to default, because the firm has a cushion of assets that it can use to pay off its loans. Hence, when firms seeking credit have high net worth, the consequences of adverse selection are less important and lenders are more willing to make loans. This concept lies behind the often-heard lament, "Only the people who don't need money can borrow it!"

Summary So far we have used the concept of adverse selection to explain seven of the eight facts about financial structure introduced earlier: The first four facts emphasize the importance of financial intermediaries and the relative unimportance of securities markets with regard to the financing of corporations; the fifth, that financial markets are among the most heavily regulated sectors of the economy; the sixth, that only large, well-established corporations have access to securities markets; and the seventh, that collateral is an important feature of debt contracts. In the next section, we will see that the other asymmetric information concept, moral hazard, provides additional reasons for the importance of financial intermediaries and the relative unimportance of securities

markets as related to the financing of corporations, the prevalence of government regulation, and the importance of collateral in debt contracts. In addition, the concept of moral hazard can be used to explain our final fact (fact 8): why debt contracts are complicated legal documents that place substantial restrictions on the behaviour of the borrower.

8.5 HOW MORAL HAZARD AFFECTS THE CHOICE BETWEEN DEBT AND EQUITY CONTRACTS

LO 8.5 Recognize the principal–agent problem arising from moral hazard in equity contracts and summarize the methods for reducing it.

Moral hazard is the asymmetric information problem that occurs after a financial transaction takes place, when the seller of a security may have incentives to hide information and engage in activities that are undesirable for the purchaser of the security. Moral hazard has important consequences for whether a firm finds it easier to raise funds with debt than with equity contracts.

Moral Hazard in Equity Contracts: The Principal–Agent Problem

Equity contracts, such as common stock, are claims to a share in the profits and assets of a business. Equity contracts are subject to a particular type of moral hazard called the **principal–agent problem**. When managers own only a small fraction of the firm they work for, the stockholders who own most of the firm's equity (called the *principals*) are not the same people as the managers of the firm. Thus the managers are the *agents* of the owners. This separation of ownership and control involves moral hazard, in that the managers in control (the agents) may act in their own interest rather than in the interest of the stockholder-owners (the principals) because the managers have less incentive to maximize profits than the stockholder-owners do.

To understand the principal–agent problem more fully, suppose that your friend Steve asks you to become a silent partner in his ice-cream store. The store setup requires an initial investment of $10 000, and Steve has only $1000. So you purchase an equity stake (stock shares) for $9000, which entitles you to 90% of the ownership of the firm, while Steve owns only 10%. If Steve works hard to make tasty ice cream, keeps the store clean, smiles at all the customers, and hustles to wait on tables quickly, after all expenses (including Steve's salary) have been paid, the store will make $50 000 in profits per year, of which Steve will receive 10% ($5000) and you will receive 90% ($45 000).

But if Steve doesn't provide quick and friendly service to his customers, uses the $50 000 in income to buy artwork for his office, and even sneaks off to the beach while he should be at the store, the store will not earn any profit. Steve can earn the additional $5000 (his 10% share of the profits) over his salary only if he works hard and forgoes unproductive investments (such as art for his office). Steve might decide that the extra $5000 just isn't enough to make him expend the effort to be a good manager. If Steve feels this way, he does not have enough incentive to be a good manager and will end up with a beautiful office, a good tan, and a store that doesn't show any profits. Because the store won't show any profits, Steve's decision not to act in your interest will cost you $45 000 (your 90% of the profits had he chosen to be a good manager instead).

The moral hazard arising from the principal–agent problem might be even worse if Steve is not totally honest. Because his ice-cream store is a cash business, Steve has the incentive to pocket $50 000 in cash and tell you that the profits were zero. He now gets a return of $50 000, and you get nothing.

Further proof that the principal–agent problem created by equity contracts can be severe is provided by past scandals involving corporations such as Enron and Tyco International, in which managers were found to have diverted corporate funds for their own personal use. In addition to pursuing personal benefits, managers might also pursue corporate strategies (such as the acquisition of other firms) that enhance their personal power but do not increase the corporation's profitability.

The principal–agent problem would not arise if the owners of a firm had complete information about what the managers were up to and could prevent wasteful expenditures or fraud. The principal–agent problem, which is an example of moral hazard, arises only because a manager, such as Steve, has more information about his activities than the stockholder does—that is, information is asymmetric. The principal–agent problem also would not occur if Steve alone owned the store, and ownership and control were not separate. If this were the case, Steve's hard work and avoidance of unproductive investments would yield him a profit (and extra income) of $50 000, an amount that would make it worth his while to be a good manager.

Tools to Help Solve the Principal–Agent Problem

Production of Information: Monitoring You have seen that the principal–agent problem arises because managers have more information about their activities and actual profits than stockholders do. One way for stockholders to reduce this moral hazard problem is for them to engage in a particular type of information production: monitoring the firm's activities by auditing the firm frequently and checking on what the management is doing. The problem is that the monitoring process can be expensive in terms of time and money, as reflected in the name economists give it: **costly state verification**. Costly state verification makes the equity contract less desirable and explains in part why equity is not a more important element in our financial structure.

As with adverse selection, the free-rider problem decreases the amount of private information production that would reduce the moral hazard (principal–agent) problem. In this example, the free-rider problem decreases monitoring. If you know that other stockholders are paying to monitor the activities of the company you hold shares in, you can take a free ride on their activities. Then you can use the money you save by not engaging in monitoring to vacation on a Caribbean island. If you can do this, though, so can other stockholders. Perhaps all the stockholders will go to the islands, and no one will spend any resources on monitoring the firm. The moral hazard problem for shares of common stock will then be severe, making it hard for firms to issue them to raise capital (providing an additional explanation for fact 1).

Government Regulation to Increase Information As with adverse selection, the government has an incentive to try to reduce the moral hazard problem created by asymmetric information, which provides another reason why the financial system is so heavily regulated (fact 5). Governments everywhere have laws to force firms to adhere to standard accounting principles that make profit verification easier. They also pass laws to impose stiff criminal penalties on people who commit the fraud of hiding and stealing profits. However, these measures can be only partly effective.

Catching this kind of fraud is not easy; fraudulent managers have the incentive to make it very hard for government agencies to find or prove fraud.

Financial Intermediation Financial intermediaries have the ability to avoid the free-rider problem in the face of moral hazard, and this is another reason why indirect finance is so important (fact 3). Two types of financial intermediaries are specifically designed to reduce the moral hazard arising from the principal–agent problem: venture-capital firms and private-equity firms. The **venture-capital firm** pools the resources of its partners and uses the funds to help budding entrepreneurs start *new* businesses. The **private-equity firm** has a similar structure to that of a venture-capital firm, but instead of investing in new businesses, it buys the shares of *existing* corporations. In exchange for supplying capital, venture-capital and private-equity firms receive an equity share in the business. Because verification of earnings and profits is so important in eliminating moral hazard, venture-capital and private-equity firms usually insist on having several of their own people participate as members of the managing body—the board of directors—of the new business so that they can keep a close watch on the new firm's activities. A key feature of venture-capital and private-equity firms is that the equity in the business they have bought is private—that is, not marketable to anyone *except* the venture-capital and private-equity firms. Thus other investors are unable to take a free ride on these firm's verification activities. As a result of this arrangement, venture-capital and private-equity firms are able to garner the full benefits of their verification activities and are given the appropriate incentives to reduce the moral hazard problem. Private-equity firms have played an important role in improving the efficiency of many corporations, while venture-capital firms have been important in the development of the high-tech sector in the United States. Both of these financial intermediaries have promoted economic growth and increased international competitiveness.

Debt Contracts Moral hazard arises with an equity contract, which is a claim on profits in all situations, whether the firm is making or losing money. If a contract could be structured so that moral hazard would exist only in certain situations, the need to monitor managers would be reduced, and the contract would be more attractive than the equity contract. The debt contract has exactly these attributes because it is a contractual agreement by the borrower to pay the lender *fixed* dollar amounts at periodic intervals. When a firm has high profits, the lender receives the contractual payments and does not need to know the exact profits of the firm. If the managers are hiding profits or pursuing activities that are personally beneficial but don't increase the firm's profitability, the lender doesn't care, as long as these activities do not interfere with the ability of the firm to make its debt payments on time. Only when the firm cannot meet its debt payments, thereby putting itself in a state of default, does the lender need to verify the state of the firm's profits. Only in this situation do lenders involved in debt contracts need to act more like equity holders; to get their fair share, they now must know how much income the firm has.

The less frequent need to monitor the firm, and thus the lower cost of state verification, helps explain why debt contracts are used more frequently than equity contracts to raise capital. The concept of moral hazard therefore helps explain fact 1, which states that stocks are not the most important source of financing for businesses.[3]

[3]Another factor that encourages the use of debt contracts rather than equity contracts in Canada is our tax code. Debt interest payments are a deductible expense for Canadian firms, whereas dividend payments to equity shareholders are not.

8.6 HOW MORAL HAZARD INFLUENCES FINANCIAL STRUCTURE IN DEBT MARKETS

LO 8.6 Summarize the methods used to reduce moral hazard in debt contracts.

Even with the advantages just described, debt contracts are still subject to moral hazard. Because a debt contract requires the borrowers to pay out a fixed amount and lets them keep any profits above this amount, the borrowers have an incentive to take on investment projects that are riskier than the lenders would like.

For example, suppose you are concerned about the problem of verifying the profits of Steve's ice-cream store, and so you decide not to become an equity partner. Instead, you lend Steve the $9000 he needs to set up his business, and you and he sign a debt contract that pays you an interest rate of 10%. As far as you are concerned, this is a surefire investment, because the demand for ice cream is strong and steady in your neighbourhood. However, once you give Steve the funds, he might use them for purposes other than what he proposed to you. Instead of opening up the ice-cream store, Steve might use your $9000 loan to invest in chemical research equipment because he thinks he has a 1-in-10 chance of inventing a diet ice cream that tastes every bit as good as the premium brands but has no fat or calories.

Obviously, this is a very risky investment, but if Steve is successful, he will become a multimillionaire. He has a strong incentive to undertake the riskier investment with your money because the gains to him would be so large if he succeeded. You would clearly be very unhappy if Steve used your loan for the riskier investment, because if he were unsuccessful, which is highly likely, you would lose most, if not all, of the money you gave him. And if he were successful, you wouldn't share in his success—you would still get only a 10% return on the loan because the principal and interest payments are fixed. Because of the potential moral hazard (that Steve might use your money to finance a very risky venture), you would probably not make the loan to Steve, even though an ice-cream store in the neighbourhood is a good investment that would provide benefits to everyone.

Tools to Help Solve Moral Hazard in Debt Contracts

Net Worth and Collateral When borrowers have more at stake because their net worth (the difference between their assets and their liabilities) is high or the collateral they have pledged to the lender is valuable, the risk of moral hazard—the temptation to act in a manner that lenders find objectionable—is greatly reduced because the borrowers themselves have a lot to lose. In other words, if borrowers have more "skin in the game" because they have higher net worth or pledge collateral, they are likely to take less risk at the lender's expense. Let's return to Steve and his ice-cream business. Suppose that the cost of setting up either the ice-cream store or the research equipment is $100 000 instead of $10 000. Now Steve needs to invest $91 000 (instead of $1000) of his own money in the business, in addition to the $9000 supplied by your loan. If Steve is unsuccessful in inventing the no-calorie, nonfat ice cream, he has a lot to lose—the $91 000 of net worth (the $100 000 in assets minus the $9000 loan from you). He will think twice about undertaking the riskier investment and is more likely to invest in the ice-cream store, which is more of a sure thing. Thus, when Steve has more of his own money (net worth) invested in the business, and hence more skin in the game, you are more likely to

make him the loan. Similarly, if you have pledged your house as collateral, you are less likely to go to Las Vegas and gamble away your earnings that month, because you might not be able to make your mortgage payments and therefore might lose your house.

One way of describing the solution that high net worth and collateral provide to the moral hazard problem is to say that it makes the debt contract **incentive-compatible**; that is, it aligns the incentives of the borrower with those of the lender. The greater the borrower's net worth and collateral pledged, the greater the borrower's incentive to behave in the way that the lender expects and desires, the smaller the moral hazard problem in the debt contract, and the easier it is for the firm or household to borrow. Conversely, when the borrower's net worth and collateral are lower, the moral hazard problem is greater, making it harder to borrow.

Monitoring and Enforcement of Restrictive Covenants As the example of Steve and his ice-cream store shows, if you could make sure that Steve doesn't invest in anything riskier than the ice-cream store, it would be worth your while to make him the loan. You can ensure that Steve uses your money for the purpose *you* expect by writing provisions (restrictive covenants) into the debt contract that restrict his firm's activities. By monitoring Steve's activities to check whether he is complying with the restrictive covenants and by enforcing the covenants if he is not, you can make sure that he will not take on risks at your expense. Restrictive covenants are directed at reducing moral hazard either by ruling out undesirable behaviour or by encouraging desirable behaviour. Four types of restrictive covenants achieve this objective:

1. ***Covenants to discourage undesirable behaviour.*** Covenants can be designed to lower moral hazard by keeping the borrower from engaging in the undesirable behaviour of undertaking risky investment projects. Some covenants mandate that a loan be used only to finance specific activities, such as the purchase of particular equipment or inventories. Others restrict the borrowing firm from engaging in certain risky business activities, such as purchasing other businesses.
2. ***Covenants to encourage desirable behaviour.*** Restrictive covenants can encourage the borrower to engage in desirable activities that make it more likely that the loan will be paid off. One such restrictive covenant requires the breadwinner in a household to carry life insurance that will pay off the mortgage upon that person's death. For businesses, restrictive covenants of this type focus on encouraging the borrowing firm to keep its net worth high because higher borrower net worth reduces moral hazard and makes it less likely that the lender will suffer losses. These restrictive covenants typically specify that the borrowing firm must maintain minimum holdings of certain assets relative to the firm's size.
3. ***Covenants to keep collateral valuable.*** Because collateral is an important protection for the lender, restrictive covenants can encourage the borrower to keep the collateral in good condition and make sure that it stays in the possession of the borrower. This is the type of covenant ordinary people encounter most often. Automobile loan contracts, for example, require the car owner to maintain a minimum amount of collision and theft insurance and prevent the sale of the car unless the loan is paid off. Similarly, the recipient of a home mortgage must have adequate insurance on the home and must pay off the mortgage when the property is sold.
4. ***Covenants to provide information.*** Restrictive covenants also require a borrowing firm to provide information about its activities periodically, in the form of quarterly accounting and income reports, thereby making it easier for the lender to monitor

the firm and reduce moral hazard. This type of covenant may also stipulate that the lender has the right to audit and inspect the firm's books at any time.

We now see why debt contracts are often complicated legal documents with numerous restrictions on the borrower's behaviour (fact 8): Debt contracts require complicated restrictive covenants to lower moral hazard.

Financial Intermediation Although restrictive covenants help reduce the moral hazard problem, they do not eliminate it completely. It is almost impossible to write covenants that rule out *every* risky activity. Furthermore, borrowers may be clever enough to find loopholes in restrictive covenants that make them ineffective.

Another problem with restrictive covenants is that they must be monitored and enforced. A restrictive covenant is meaningless if the borrower can violate it knowing that the lender won't check up or is unwilling to pay for legal recourse. Because monitoring and enforcement of restrictive covenants are costly, the free-rider problem arises in the debt securities (bond) market just as it does in the stock market. If you know that other bondholders are monitoring and enforcing the restrictive covenants, you can free ride on their monitoring and enforcement. But other bondholders can do the same thing, so the likely outcome is that not enough resources will be devoted to monitoring and enforcing the restrictive covenants. Moral hazard therefore continues to be a severe problem for marketable debt.

As we have seen before, financial intermediaries—particularly banks—can avoid the free-rider problem by primarily making *private* loans. Private loans are not traded, so no one else can free ride on the intermediary's monitoring and enforcement of the restrictive covenants. The intermediary making private loans thus receives the full benefits of monitoring and enforcement and will work to shrink the moral hazard problem inherent in debt contracts. The concept of moral hazard has provided us with additional reasons why financial intermediaries play a more important role in channelling funds from savers to borrowers than marketable securities do, as described in facts 3 and 4.

Summary

Asymmetric information in financial markets leads to adverse selection and moral hazard problems that interfere with the efficient functioning of those markets. Tools to help solve these problems involve the private production and sale of information, government regulation to increase information in financial markets, the importance of collateral and net worth to debt contracts, and the use of monitoring and restrictive covenants. Our analysis suggests that the existence of the free-rider problem for traded securities such as stocks and bonds indicates that financial intermediaries—particularly banks—should play a greater role than securities markets in financing the activities of businesses. Economic analysis of the consequences of adverse selection and moral hazard has helped explain the basic features of our financial system, including the eight facts about our financial structure outlined at the beginning of this chapter.

As a study aid, Table 8-1 summarizes the asymmetric information problems and tools that help solve them. In addition, it notes how these tools and asymmetric information problems explain the eight facts of financial structure described at the beginning of the chapter.

SUMMARY TABLE 8-1

Asymmetric Information Problems and Tools to Solve Them

Asymmetric Information Problem	Tools to Solve It	Explains Fact Number
Adverse selection	Private production and sale of information	1, 2
	Government regulation to increase information	5
	Financial intermediation	3, 4, 6
	Collateral and net worth	7
Moral hazard in equity contracts (principal–agent problem)	Production of information: monitoring	1
	Government regulation to increase information	5
	Financial intermediation	3
	Debt contracts	1
Moral hazard in debt contracts	Collateral and net worth	6, 7
	Monitoring and enforcement of restrictive covenants	8
	Financial intermediation	3, 4

Note: List of facts:
1. Stocks are not the most important source of external financing.
2. Marketable securities are not the primary source of financing.
3. Indirect finance is more important than direct finance.
4. Banks are the most important source of external funds.
5. The financial system is heavily regulated.
6. Only large, well-established firms have access to securities markets.
7. Collateral is prevalent in debt contracts.
8. Debt contracts have numerous restrictive covenants.

APPLICATION Financial Development and Economic Growth

Recent research has found that an important reason why many developing countries and ex-communist countries like Russia (which are referred to as *transition countries*) experience very low rates of economic growth is that their financial systems are underdeveloped (a situation referred to as *financial repression*).[4] The economic analysis of financial structure helps explain how an underdeveloped financial system leads to a low state of economic development and economic growth.

The financial systems of developing and transition countries face several difficulties that keep them from operating efficiently. As we have seen, two important tools used to

[4]See World Bank, *Finance for Growth: Policy Choices in a Volatile World* (Washington, DC: World Bank and Oxford University Press, 2001), and Frederic S. Mishkin, *The Next Great Globalization: How Disadvantaged Nations Can Harness Their Financial Systems to Get Rich* (Princeton, NJ: Princeton University Press, 2006) for a survey of the literature linking economic growth with financial development and a list of additional references.

FYI The Tyranny of Collateral

To use property, such as land or capital, as collateral, a person must legally own it. Unfortunately, as Hernando De Soto documents in his book *The Mystery of Capital*, it is extremely expensive and time-consuming for the poor in developing countries to make their ownership of property legal. Obtaining legal title to a dwelling on urban land in the Philippines, for example, involved 168 bureaucratic steps and 53 public and private agencies, and the process took anywhere from 13 to 25 years. Obtaining legal title to desert land in Egypt took 77 steps, 31 public and private agencies, and anywhere from 5 to 14 years. To legally buy government land in Haiti, an ordinary citizen had to go through 176 steps over 19 years. These legal barriers do not mean that the poor do not invest: They still build houses and buy equipment, even if they don't have legal title to these assets. By De Soto's calculations, the "total value of the real estate held but not legally owned by the poor of the Third World and former communist nations is at least $9.3 trillion."*

Without legal title, however, none of this property can be used as collateral to borrow funds, a requirement for most lenders. Even when people have legal title to their property, the legal system in most developing countries is so inefficient that collateral does not mean much. Typically, creditors must first sue the defaulting debtor for payment, which takes several years, and then, upon obtaining a favourable judgment, the creditor has to sue again to obtain title to the collateral. This process often takes in excess of five years. By the time the lender acquires the collateral, it is likely to have been neglected or stolen and thus has little value. In addition, governments often block lenders from foreclosing on borrowers in politically powerful sectors of a society, such as agriculture.

When the financial system is unable to use collateral effectively, the adverse selection problem worsens because the lender needs even more information about the quality of the borrower in order to distinguish a good loan from a bad one. Little lending will take place, especially in transactions that involve collateral, such as mortgages. In Peru, for example, the value of mortgage loans relative to the size of the economy is less than 1/20 that in the United States.

The poor in developing countries have an even harder time obtaining loans because it is too costly for them to get title to their property, and therefore they have no collateral to offer, resulting in what Raghuram Rajan, the governor of India's central bank, and Luigi Zingales of the University of Chicago refer to as the "tyranny of collateral."** Even when poor people have a good idea for a business and are willing to work hard, they cannot get the funds to finance the business, making it difficult for them to escape poverty.

* Hernando De Soto, *The Mystery of Capital: Why Capitalism Triumphs in the West and Fails Everywhere Else* (New York: Basic Books, 2000), 35.

** Raghuram Rajan and Luigi Zingales, *Saving Capitalism from the Capitalists: Unleashing the Power of Financial Markets to Create Wealth and Spread Opportunity* (New York: Crown Business, 2003).

help solve adverse selection and moral hazard problems in credit markets are collateral and restrictive covenants. In many developing countries, the system of property rights (the rule of law, constraints on government expropriation, and the absence of corruption) functions poorly, making it hard for these two tools to function effectively.

As discussed in the FYI box "The Tyranny of Collateral," the weak systems of property rights common in developing and transition countries impede the use of collateral, making adverse selection problems worse because the lender will need even more information about the quality of the borrower in order to screen out a good loan from a bad one. As a result, it is harder for lenders to channel funds to the borrowers with the most productive investment opportunities, ultimately leading to less productive investment and hence a slower-growing economy.

Similarly, a poorly developed or corrupt legal system makes it extremely difficult for lenders to enforce restrictive covenants. Thus their ability to reduce moral hazard on the part of borrowers is severely limited, and so they are less willing to lend. Again, the outcome is less productive investment and a lower growth rate for the economy. The importance of an effective legal system in promoting economic growth suggests that lawyers play a more positive role in the economy than we give them credit for.

Governments in developing and transition countries often use their financial systems to direct credit to themselves or to favoured sectors of the economy by setting interest rates at artificially low levels for certain types of loans, by creating development finance institutions to make specific types of loans, or by directing existing institutions to lend to certain entities. As we have seen, private institutions have an incentive to solve adverse selection and moral hazard problems and to lend to borrowers with the most productive investment opportunities. Governments have less incentive to do so because they are not driven by the profit motive and thus their directed credit programs may not channel funds to sectors that will produce high growth for the economy. The likely outcome is, again, less efficient investment and slower economic growth.

In addition, banks in many developing and transition countries are owned by their governments. Again, because of the absence of the profit motive, these **state-owned banks** have little incentive to allocate their capital to the most productive uses. Not surprisingly, the primary loan customer of these state-owned banks is often the government itself, which does not always use the funds wisely.

We have seen that government regulation can increase the amount of information available in financial markets and thus help them work more efficiently. Many developing and transition countries have an underdeveloped regulatory apparatus that retards the provision of adequate information to the marketplace. For example, these countries often have weak accounting standards, making it very hard to ascertain the quality of a borrower's balance sheet. As a result, asymmetric information problems are more severe, and the financial system is severely hampered in channelling funds to the most productive uses.

The institutional environment of a poor legal system, weak accounting standards, inadequate government regulation, and government intervention through directed credit programs and state ownership of banks all help explain why many countries stay poor while others grow rich. ◆

APPLICATION Is China a Counterexample to the Importance of Financial Development?

Although China appears to be on its way to becoming an economic powerhouse, its financial development remains in the early stages. The country's legal system is weak, so that financial contracts are difficult to enforce, and accounting standards are lax, making high-quality information about creditors hard to obtain. Regulation of the banking system is still in its formative stages, and the banking sector is dominated by large, state-owned banks. Yet the Chinese economy has enjoyed one of the highest growth rates in the world over the past 30 years. How has China been able to grow so rapidly, given its low level of financial development?

As noted above, China is in an early state of financial development, with a per capita income that is around $20 000, less than a third of per capita income in the

United States. With an extremely high savings rate, averaging around 40% over the past three decades, the country has been able to rapidly build up its capital stock and shift a massive pool of underutilized labour from the subsistence-agriculture sector into higher-productivity activities that use capital. Even though available savings have not always been allocated to their most productive uses, the huge increase in capital, combined with the gains in productivity achieved by moving labour out of low-productivity, subsistence agriculture, have been enough to produce high growth.

As China gets richer, however, this strategy is unlikely to continue to work. The Soviet Union provides a graphic example. In the 1950s and 1960s, the Soviet Union shared many characteristics with modern-day China: high growth fuelled by a high savings rate, a massive buildup of capital, and shifts of a large pool of underutilized labour from subsistence agriculture to manufacturing. During this high-growth phase, however, the Soviet Union was unable to develop the institutions needed to allocate capital efficiently. As a result, once the pool of subsistence labourers was used up, the Soviet Union's growth slowed dramatically and it was unable to keep up with the Western economies. Today, no one considers the Soviet Union an economic success story, and its inability to develop the institutions necessary to sustain financial development and growth is an important reason for the demise of this superpower.

To move into the next stage of development, China will need to allocate its capital more efficiently, which requires that it improve its financial system. The Chinese leadership is well aware of this challenge; the government has announced that state-owned banks are being put on the path to privatization. In addition, the government is engaged in legal reform aimed at making financial contracts more enforceable. New bankruptcy law is being developed so that lenders will have the ability to take over the assets of firms that default on their loan contracts. Whether the Chinese government will succeed in developing a first-rate financial system, thereby enabling China to join the ranks of developed countries, is a big question mark. ◆

SUMMARY

1. Eight basic facts summarize financial structure throughout the world. The first four emphasize the importance of financial intermediaries and the relative unimportance of securities markets with regard to the financing of corporations; the fifth recognizes that financial markets are among the most heavily regulated sectors of the economy; the sixth states that only large, well-established corporations have access to securities markets; the seventh indicates that collateral is an important feature of debt contracts; and the eighth presents debt contracts as complicated legal documents that place substantial restrictions on the behaviour of the borrower.
2. Transaction costs freeze many small savers and borrowers out of direct involvement with financial markets. Financial intermediaries can take advantage of economies of scale and are better able to develop expertise, leading to lower transaction costs and thus enabling their savers and borrowers to benefit from the existence of financial markets.
3. Asymmetric information results in two problems: adverse selection, which occurs before the transaction takes place, and moral hazard, which occurs after the transaction has taken place. Adverse selection refers to the fact that bad credit risks are the ones most likely to seek loans, and moral hazard refers to the risk of the borrower engaging in activities that are undesirable from the lender's point of view.
4. Adverse selection interferes with the efficient functioning of financial markets. Tools that help reduce the adverse selection problem include private production and sale of information, government regulation to increase information, financial intermediation, and

collateral and net worth. The free-rider problem occurs when people who do not pay for information take advantage of information that other people have paid for. This problem explains why financial intermediaries, particularly banks, play a more important role in financing the activities of businesses than do securities markets.

5. Moral hazard in equity contracts, known as the principal–agent problem, occurs because managers (the agents) have less incentive to maximize profits than stockholders (the principals). The principal–agent problem explains why debt contracts are so much more prevalent in financial markets than equity contracts. Tools that help reduce the principal–agent problem include monitoring, government regulation to increase information, and financial intermediation.
6. Tools that are used to reduce the moral hazard problem in debt contracts include net worth, monitoring and enforcement of restrictive covenants, and financial intermediaries.

MyLab Economics Make the grade with MyLab Economics: Select end-of-chapter questions can be found on MyLab Economics. You can practise them as often as you want, and most feature step-by-step guided instructions to help you find the right answer.

QUESTIONS

1. For each of the following countries, identify the single most important (largest) and least important (smallest) source of external funding: Canada, United States, Germany, Japan. Comment on the similarities and differences among the countries' funding sources.
2. How can economies of scale help explain the existence of financial intermediaries?
3. Explain why dating can be considered a method to solve the adverse selection problem.
4. Why are financial intermediaries willing to engage in information collection activities when investors in financial instruments may be unwilling to do so?
5. Suppose you go to your local bank, intending to buy a certificate of deposit with your savings. Explain why you would prefer this to offering a loan, at an interest rate that is higher than the rate the bank pays on certificates of deposit (but lower than the rate the bank charges for car loans), to the next individual who enters the bank and applies for a car loan.
6. Suppose you are applying for a mortgage loan. The loan officer tells you that if you get the loan, the bank will keep the house title until you pay back the loan. Which problem of asymmetric information is the bank trying to solve?
7. Suppose you have data about two groups of countries, one with efficient legal systems and the other with slow, costly, and inefficient legal systems. Which group of countries would you expect to exhibit higher living standards?
8. Which relationship would you expect to exist between measures of corruption and living standards at the country level? Explain by which channel corruption might affect living standards.
9. Would you be more willing to lend to a friend if they had put all of their life savings into their business than you would be if they had not done so? Why?
10. What steps can the government take to reduce asymmetric information problems and help the financial system function more smoothly and efficiently?
11. How can asymmetric information problems lead to a bank panic?
12. In December 2001, Argentina announced it would not honour its sovereign (government-issued) debt. Many investors were left holding Argentinean bonds priced at a fraction of their previous value. A few years later, Argentina announced it would pay back 25% of the face value of its debt. Comment on the effects of information asymmetries on government bond markets. Do

you think investors are currently willing to buy bonds issued by the government of Argentina?

13. How does the free-rider problem aggravate adverse selection and moral hazard problems in financial markets?
14. Suppose that in a given bond market, there is currently no information that can help potential bond buyers to distinguish between bonds. Which bond issuers have an incentive to disclose information about their companies? Explain why.
15. How do standardized accounting principles help financial markets work more efficiently?
16. Which problem of asymmetric information are prospective employers trying to solve when they ask applicants to go through a job interview? Is that the end of the information asymmetry?
17. How can the existence of asymmetric information provide a rationale for government regulation of financial markets?
18. "The more collateral there is backing a loan, the less the lender has to worry about adverse selection." Is this statement true, false, or uncertain? Explain your answer.
19. Explain how the separation of ownership and control in Canadian corporations might lead to poor management.
20. Many policymakers in developing countries have proposed the implementation of a system of deposit insurance similar to the system that exists in Canada. Explain why this might create more problems than solutions in the financial system of a developing country.
21. Gustavo is a young doctor who lives in a country with a relatively inefficient legal and financial system. When Gustavo applied for a mortgage, he found that banks usually required collateral for up to 300% of the amount of the loan. Explain why banks might require that much collateral in such a financial system. Comment on the consequences of such a system for economic growth.

APPLIED PROBLEMS

For Problems 22–25, use the fact that the expected value of an event is a probability weighted average, the sum of each possible outcome multiplied by the probability of the event occurring.

22. You are in the market for a used car and decide to visit a used-car dealership. You know that the Blue Book value of the car you are looking at is between $20 000 and $24 000. If you believe the dealer knows *as much* about the car as you do, how much are you willing to pay? Why? Assume that you care about only the expected value of the car you will buy and that the car values are symmetrically distributed.
23. Refer to Problem 22. Now you believe the dealer knows *more* about the car than you do. How much are you willing to pay? Why? How can this asymmetric information problem be resolved in a competitive market?
24. You wish to hire Ron to manage your Calgary operations. The profits from the operations depend partially on how hard Ron works, as follows.

Profit Probabilities

	Profit = $10 000	Profit = $50 000
Lazy	60%	40%
Hard worker	20%	80%

If Ron is lazy, he will surf the Internet all day, and he views this as a zero-cost opportunity. However, Ron views working hard as a "personal cost" valued at $1000. What fixed percentage of the profits should you offer Ron? Assume Ron cares only about his expected payment less any "personal cost."

25. You own a house worth $400 000 that is located on a river. If the river floods moderately, the house will be completely destroyed. Moderate flooding happens about once every 50 years. If you build a seawall, the river would have to flood heavily to destroy your house, and such heavy flooding happens only about once every 200 years. What would be the annual premium for a flood insurance policy that offers full insurance? For a policy that pays only 75% of the home value, what are your expected costs with and without a seawall? Do the different policies provide an incentive to be safer (i.e., to build the seawall)?

DATA ANALYSIS PROBLEMS

1. Go to the Statistics Canada CANSIM database and find quarterly series (from 2000Q1 to 2017Q2) for the net new issues of securities by financial and non-financial corporations, series V122334 (total bonds) and V122335 (total stocks).
    a. Compare the bonds series to the stocks series. Which is a more important source of finance, stocks or bonds?
    b. What is the correlation coefficient between the two series? Has the correlation changed in the aftermath of the global financial crisis (after 2008)?
2. **Real-Time Data Analysis** Go to the St. Louis Federal Reserve FRED database and find data on net worth of households (BOGZ1FL192090005Q) and the net percentage of domestic banks tightening standards for auto loans (STDSAUTO). Adjust the *units* setting for the net worth indicator to "Percent Change from Year Ago," and download the data into a spreadsheet.
    a. Calculate the average, over the most recent four quarters and the four quarters prior to that, for the bank standards indicator and the "percent change in net worth" indicator. Do these averages behave as you would expect?
    b. Use the Data Analysis tool in Excel to calculate the correlation coefficient for the two data series from 2011:Q2 to the most recent quarter of data available. What can you conclude about the relationship between the net worth of households and bank auto lending standards? Is this result consistent with efforts to reduce asymmetric information?

9 Economic Analysis of Financial Regulation

Learning Objectives

9.1 Identify the reasons for and the forms of a government safety net in financial markets.

9.2 List and summarize the types of financial regulation and how each reduces asymmetric information problems.

9.3 Examine how the Canada Deposit Insurance Corporation (CDIC) provides deposit insurance in Canadian banks and savings institutions.

Preview

As we have seen in the previous chapters, the financial system is among the most heavily regulated sectors of the economy, and banks are among the most heavily regulated of financial institutions. In this chapter, we develop an economic analysis of why regulation of the financial system takes the form it does.

Unfortunately, the regulatory process may not always work very well, as evidenced by the global financial crisis of 2007–2009. Here we also use our economic analysis of financial regulation to explain the worldwide crises in banking and to consider how the regulatory system can be reformed to prevent future such disasters.

9.1 ASYMMETRIC INFORMATION AND THE GOVERNMENT SAFETY NET

LO 9.1 Identify the reasons for and forms of a government safety net in financial markets.

In earlier chapters, we have seen how asymmetric information—the fact that different parties in a financial contract do not have the same information—leads to adverse selection and moral hazard problems, which have an important impact on our financial system. The concepts of asymmetric information, adverse selection, and moral hazard are especially useful in understanding why governments pursue financial regulation.

Government Safety Net

As we saw in Chapter 8, financial intermediaries, like banks, are particularly well suited to solving adverse selection and moral hazard problems because they make private loans that help avoid the free-rider problem. However, this solution to the free-rider problem creates another asymmetric information problem because depositors lack information about the quality of these private loans. This asymmetric information problem leads to other problems that interfere with the proper functioning of the financial system.

Bank Panics and the Need for Deposit Insurance Before the Canada Deposit Insurance Corporation (CDIC) started operations in 1967, a **bank failure** (in which a bank is unable to meet its obligations to pay its depositors and other creditors and so must go out of business) meant that depositors would have to wait until

the bank was liquidated (until its assets had been turned into cash) to get their deposit funds; at that time, they would be paid only a fraction of the value of their deposits. Because they couldn't know if bank managers were taking on too much risk or were outright crooks, depositors would be reluctant to put money in banks, thus making financial institutions less viable. Further, depositors' lack of information about the quality of bank assets could lead to a **bank panic**, in which many banks fail simultaneously. Because the simultaneous failure of many banks leads to a sharp decline in bank lending, bank panics have serious, harmful consequences for the economy.

To understand why bank panics occur, consider the following situation: Deposit insurance does not exist, and an adverse shock hits the economy. As a result of the shock, 5% of banks have such large losses on loans that they become insolvent (have a negative net worth and so are bankrupt). Because of asymmetric information, depositors are unable to tell whether their bank is a "good" bank or one of the 5% that are insolvent. Depositors at bad *and* good banks recognize that they may not get back 100 cents on the dollar for their deposits and therefore are inclined to withdraw them. Indeed, because banks operate on a "sequential service constraint" (a first-come, first-served basis), depositors have a very strong incentive to be the first to show up at the bank, because if they are last in line, the bank may have paid out all its funds, and they will get nothing. The incentive to "run" to the bank to be first is why withdrawals when there is fear about the health of a bank is described as a "bank run." Uncertainty about the health of the banking system in general can lead to runs on both good and bad banks, and the failure of one bank can hasten the failure of others (referred to as the *contagion effect*). If nothing is done to restore the public's confidence, a bank panic can ensue.

A government safety net for depositors can short-circuit runs on banks and bank panics, and by providing protection for the depositor, it can overcome depositors' reluctance to put funds in the banking system. One form of the safety net is deposit insurance, such as that provided by the CDIC. The CDIC guarantees that current depositors will be paid off in full on the first $100 000 they have deposited in the bank if the bank fails; in the United States, the Federal Deposit Insurance Corporation (FDIC) guarantees deposits up to $250 000. With fully insured deposits, depositors don't need to run to the bank to make withdrawals—even if they are worried about the bank's health—because their deposits will be worth 100 cents on the dollar no matter what.

The CDIC uses two primary methods to handle a failed bank. In the first, called the *payoff method*, the CDIC allows the bank to fail and pays off deposits up to the $100 000 insurance limit (with funds acquired from the insurance premiums paid by the banks that have bought CDIC insurance). After the bank has been liquidated, the CDIC lines up with other creditors of the bank and is paid its share of the proceeds from the liquidated assets. Typically, when the payoff method is used, account holders with deposits in excess of the $100 000 limit get back more than 90 cents on the dollar, although the process can take several years to complete.

In the second method, called the *purchase-and-assumption method*, the CDIC reorganizes the bank, typically by finding a willing merger partner who assumes (takes over) all the failed bank's liabilities so that no depositor or other creditor loses a penny. The CDIC often sweetens the pot for the merger partner by providing it with subsidized loans or by buying some of the failed bank's weaker loans. The net effect of the purchase-and-assumption method is that the CDIC has guaranteed *all* liabilities and deposits, not just those under the $100 000 limit. The purchase-and-assumption method is typically more costly for the CDIC than the payoff method but nevertheless is the CDIC's more common procedure for dealing with a failed bank.

Global **The Spread of Government Deposit Insurance Throughout the World: Is It a Good Thing?**

Government deposit insurance has taken off throughout the world because of growing concern about the health of banking systems, particularly after the increasing number of banking crises in recent years (documented in the online appendices to this chapter). Has this spread of deposit insurance been a good thing? Has it helped improve the performance of the financial system and prevent banking crises?

The answer seems to be "no" under many circumstances. Research at the World Bank has found that, on average, the adoption of explicit government deposit insurance is associated with less banking-sector stability and a higher incidence of banking crises.* Furthermore, on average, deposit insurance seems to retard financial development. However, these negative effects of deposit insurance occur only in countries with weak institutional environments: an absence of rule of law, ineffective regulation and supervision of the financial sector, and high corruption. This situation is exactly what might be expected because, as we will see later in this chapter, a strong institutional environment is needed to limit the moral hazard incentives for banks to engage in the excessively risky behaviour encouraged by deposit insurance. The problem is that development of a strong institutional environment may be very difficult to achieve in many emerging-market countries. We are left with the following conclusion: Adoption of deposit insurance may be exactly the wrong medicine for promoting stability and efficiency of banking systems in emerging-market countries.

*See *World Bank, Finance for Growth: Policy Choices in a Volatile World* (Oxford: World Bank and Oxford University Press, 2001).

In recent years, government deposit insurance has been growing in popularity and has spread to many countries throughout the world. Whether this trend is desirable is discussed in the Global box "The Spread of Government Deposit Insurance Throughout the World: Is It a Good Thing?"

Other Forms of the Government Safety Net

Deposit insurance is not the only form of government safety net. In other countries, governments have often stood ready to provide support to domestic banks facing runs even in the absence of explicit deposit insurance. Furthermore, banks are not the only financial intermediaries that can pose a systemic threat to the financial system. When financial institutions are very large or highly interconnected with other financial institutions or markets, their failure has the potential to bring down the entire financial system. This is exactly what happened in the United States with Bear Stearns and Lehman Brothers, two investment banks, and AIG, an insurance company, during the global financial crisis in 2008.

One way governments provide support is through lending from the central bank to troubled institutions, as the Bank of Canada did during the global financial crisis (more on this in Chapter 16). This form of support is often referred to as the "lender of last resort" role of the central bank. In other cases, funds are provided directly to troubled institutions, as was done by the government of Canada through the Canada Mortgage and Housing Corporation (CMHC), by the U.S. Treasury, and by other governments in 2008 during a particularly virulent phase of the global financial crisis. Governments also take over (nationalize) troubled institutions and guarantee that all creditors' loans will be paid in full.

Drawbacks of the Government Safety Net

Although a government safety net can help protect depositors and other creditors and can prevent, or ameliorate, financial crises, it is a mixed blessing.

Moral Hazard and the Government Safety Net The most serious drawback of the government safety net stems from moral hazard, the incentives of one party in a transaction to engage in activities detrimental to the other party. Moral hazard is an important concern in insurance arrangements in general because the existence of insurance provides increased incentives for taking risks that might result in an insurance payoff. For example, some drivers who have automobile collision insurance that has a low deductible might be more likely to drive recklessly, because if they get into an accident, the insurance company pays most of the costs for damage and repairs.

Moral hazard is a prominent concern associated with government safety nets. With a safety net, depositors and creditors know that they will not suffer losses if a financial institution fails, so they do not impose the discipline of the marketplace on these financial institutions by withdrawing funds when they suspect that the financial institution is taking on too much risk. Consequently, financial institutions with a government safety net have an incentive to take on greater risks than they otherwise would, because taxpayers will foot the bill if the bank subsequently goes belly up. Financial institutions can place the following bet: "Heads I win, tails the taxpayer loses."

Adverse Selection and the Government Safety Net A further problem with a government safety net like deposit insurance is adverse selection, the fact that the people who are most likely to produce the adverse outcome the bank is insured against (bank failure) are the same people who most want to take advantage of the insurance. For example, bad drivers are more likely than good drivers to take out automobile collision insurance with a low deductible. Because depositors and creditors protected by a government safety net have little reason to impose discipline on financial institutions, risk-loving entrepreneurs might find the financial industry a particularly attractive one—they know that they will be able to engage in highly risky activities. Even worse, because protected depositors and creditors have so little reason to monitor the financial institution's activities, without government intervention outright crooks might also find finance an attractive industry for their activities because it is easy for them to get away with fraud and embezzlement.

"Too Big to Fail" The moral hazard created by a government safety net and the desire to prevent financial failures have presented financial regulators with a particular quandary, the **too-big-to-fail problem**, in which regulators are reluctant to close down large financial institutions and impose losses on their depositors and creditors because doing so might precipitate a financial crisis.

One problem with the too-big-to-fail policy is that it increases the moral hazard incentives for big banks. If the CDIC were willing to close a bank using the payoff method, paying depositors only up to the current $100 000 limit, large depositors with more than $100 000 would suffer losses if the bank failed. Thus they would have an incentive to monitor the bank by examining the bank's activities closely and pulling their money out if the bank were taking on too much risk. To prevent such a loss of deposits, the bank would be more likely to engage in less risky activities. However, once large depositors know that a bank is too big to fail, they have no incentive to monitor the bank and pull out their deposits when it takes on too much risk: No matter what the bank does, large depositors will not suffer any losses. The result of the too-big-to-fail policy is that big financial institutions might take on even greater risks, thereby making bank failures more likely.

Similarly, the too-big-to-fail policy increases the moral hazard incentives for nonbank financial institutions that are extended a government safety net. Knowing that the

financial institution will get bailed out, creditors have little incentive to monitor the institution and pull their money out when the institution is taking on excessive risk. As a result, large or interconnected financial institutions are more likely to engage in highly risky activities, making a financial crisis more likely.

Indeed, financial institutions in the United States that were considered too big to fail—including Bear Stearns, Lehman Brothers, and AIG—did take on excessive risk in the period leading to the global financial crisis, and their subsequent collapse helped trigger the worst financial crisis since the Great Depression.

Financial Consolidation and the Government Safety Net Financial consolidation has been proceeding at a rapid pace, leading to both larger and more complex financial organizations. Financial consolidation poses two challenges to financial regulation because of the existence of the government safety net.

First, the increased size of financial institutions resulting from financial consolidation increases the too-big-to-fail problem, because there are now more large institutions whose failure would expose the financial system to systemic (system-wide) risk. Thus, more financial institutions are likely to be treated as too big to fail, and the increased moral hazard incentives for these large institutions to take on greater risk increases the fragility of the financial system.

Second, financial consolidation of banks with other financial services firms means that the government safety net may be extended to new activities, such as securities underwriting, insurance, or real estate activities, as occurred in the United States during the global financial crisis. This situation increases incentives for greater risk taking in these activities, which can also weaken the fabric of the financial system. Limiting the moral hazard incentives for the larger, more complex financial organizations that have arisen as a result of recent changes in legislation is one of the key issues facing banking regulators in the aftermath of the global financial crisis.

9.2 TYPES OF FINANCIAL REGULATION

LO 9.2 List and summarize the types of financial regulation and how each reduces asymmetric information problems.

There are eight basic types of financial regulation aimed at lessening asymmetric information problems and excessive risk taking in the financial system: (1) restrictions on asset holdings, (2) capital requirements, (3) prompt corrective action, (4) chartering and examination, (5) assessment of risk management, (6) disclosure requirements, (7) consumer protection, and (8) restrictions on competition.

Restrictions on Asset Holdings

As we have seen, the moral hazard associated with a government safety net encourages excessive risk taking on the part of financial institutions. Financial regulations that restrict asset holdings are directed at minimizing this moral hazard, which can cost the taxpayers dearly.

Even in the absence of a government safety net, financial institutions still have the incentive to take on too much risk. Risky assets may provide the financial institution with higher earnings when they pay off, but if they do not pay off and the institution

fails, depositors and creditors are left holding the bag. If depositors and creditors were able to monitor the bank easily by acquiring information on its risk-taking activities, they would be able to withdraw their funds immediately if the institution was taking on too much risk. To prevent such a loss of funds, the institution would be more likely to reduce its risk-taking activities. Unfortunately, acquiring information on an institution's activities to learn how much risk the institution is taking on can be a difficult task. Hence, most depositors and many creditors are incapable of imposing discipline that might prevent financial institutions from engaging in risky activities. A strong rationale for government regulation aimed at reducing risk taking on the part of financial institutions therefore existed even before the establishment of government safety nets like deposit insurance.

Because banks are the financial institutions most prone to panics, they are subjected to strict regulations that restrict their holding of risky assets, such as common stocks. Bank regulations also promote diversification, which reduces risk by limiting the amount of loans given in particular categories or to individual borrowers. With the extension of the government safety net during the global financial crisis and the calls for regulatory reform in its aftermath, it is likely that nonbank financial institutions may face greater restrictions on their holdings of risky assets in the future. The danger exists, however, that these restrictions may become so onerous that the efficiency of the financial system will be impaired.

Capital Requirements

Government-imposed capital requirements are another way of minimizing moral hazard at financial institutions. When a financial institution is forced to hold a large amount of equity capital, the institution has more to lose if it fails and is thus more likely to pursue less risky activities. In addition, as will be illustrated in Chapter 12, capital functions as a cushion when bad shocks occur, making it less likely that a financial institution will fail, thereby directly adding to the safety and soundness of financial institutions.

Capital requirements for banks take two forms. The first type is based on the **leverage ratio**, the amount of capital divided by the bank's total assets. To be classified as well capitalized, a bank's leverage ratio must exceed 5%; a lower leverage ratio, especially one below 3%, triggers increased regulatory restrictions on the bank.

In the past decade, regulators in Canada and the rest of the world have become increasingly worried about banks' holdings of risky assets and about the increase in banks' **off-balance-sheet activities**. Off-balance-sheet activities, which do not appear on bank balance sheets but nevertheless expose banks to risk, involve trading financial instruments and generating income from fees.

To help combat the problems of risky assets and off-balance-sheet activities, banking officials from industrialized nations agreed to set up the **Basel Committee on Bank Supervision** (so named because it meets under the auspices of the Bank for International Settlements in Basel, Switzerland), which implemented the **Basel Accord**, which deals with a second type of capital requirement, risk-based capital requirements. The Basel Accord, which requires that banks hold as capital at least 8% of their risk-weighted assets, has been adopted by more than 100 countries, including Canada and the United States. Assets are allocated into four categories, each with a different weight to reflect the degree of credit risk. The first category carries a zero weight and includes items that have little default risk, such as reserves and government securities issued by industrialized countries that are members of the Organisation for Economic Co-operation and

Development (OECD). The second category has a 20% weight and includes claims on banks in OECD countries. The third category has a weight of 50% and includes municipal bonds and residential mortgages. The fourth category has the maximum weight of 100% and includes loans to consumers and corporations. Off-balance-sheet activities are treated in a similar manner. They are assigned a credit-equivalent percentage that converts them to on-balance-sheet items to which the appropriate risk weight applies, and there are minimum capital requirements for risks in banks' trading accounts.

Over time, the limitations of the Basel Accord have become apparent, because the regulatory measure of bank risk, as stipulated by the risk weights, can differ substantially from the actual risk the bank faces. This discrepancy has resulted in **regulatory arbitrage**, a practice in which banks keep on their books assets that have the same risk-based capital requirement but are relatively risky, such as a loan to a company with a very low credit rating, while taking off their books low-risk assets, such as a loan to a company with a very high credit rating. The Basel Accord thus might lead to increased risk taking, the opposite of its intent. To address these limitations, the Basel Committee on Bank Supervision came up with a new capital accord, often referred to as Basel II. However, in the aftermath of the global financial crisis, the committee developed an even newer accord, which the media has dubbed "Basel III." These accords are described in the Global box "Where Is the Basel Accord Heading After the Global Financial Crisis?"

Prompt Corrective Action

If the amount of a financial institution's capital falls to low levels, two serious problems result. First, the bank is more likely to fail because it has a smaller capital cushion if it suffers loan losses or other asset write-downs. Second, with less capital, a financial institution has less "skin in the game" and is therefore more likely to take on excessive risks. In other words, the moral hazard problem becomes more severe, making it more likely that the institution will fail and the taxpayer will be left holding the bag. To prevent this, the CDIC adopted prompt corrective action provisions that require the CDIC to intervene earlier and more vigorously when a bank gets into trouble.

Financial Supervision: Chartering and Examination

Overseeing who operates financial institutions and how they are operated, referred to as **financial supervision** or **prudential supervision**, is an important method for reducing adverse selection and moral hazard in the financial industry. Because financial institutions can be used by crooks or overambitious entrepreneurs to engage in highly speculative activities, such undesirable people would be eager to run a financial institution. Chartering financial institutions is one method for preventing this adverse selection problem; through chartering, proposals for new institutions are screened to prevent undesirable people from controlling them.

A chartered bank obtains a charter either by an Act of Parliament or through application to the minister of finance, who has the authority to issue a charter. To obtain a charter, the people planning to organize the bank must submit an application that shows how they plan to operate the bank. In evaluating the application, the regulatory authority looks at whether the bank is likely to be sound by examining the quality of the bank's intended management, the likely earnings of the bank, and the amount of the bank's initial capital. Moreover, the chartering agency typically explores the issue of whether the community needs a new bank. Often a new bank charter would not

Global Where Is the Basel Accord Heading After the Global Financial Crisis?

Starting in June 1999, the Basel Committee on Banking Supervision released several proposals to reform the original 1988 Basel Accord. These efforts have culminated in what bank supervisors refer to as Basel II, which is based on three pillars.

1. Pillar 1 links capital requirements for large, internationally active banks more closely to actual risk of three types: market risk, credit risk, and operational risk. It does so by specifying many more categories of assets, with different risk weights, in its standardized approach. Alternatively, it allows sophisticated banks to pursue an internal ratings-based approach that permits banks to use their own models of credit risk.
2. Pillar 2 focuses on strengthening the supervisory process, particularly in assessing the quality of risk management in banking institutions and evaluating whether these institutions have adequate procedures in place for determining how much capital they need.
3. Pillar 3 focuses on improving market discipline through increased disclosure of details about a bank's credit exposures, its amount of reserves and capital, the officials who control the bank, and the effectiveness of its internal rating system.

Although Basel II made great strides toward limiting excessive risk taking by internationally active banking institutions, it greatly increased the complexity of the accord. The document describing the original Basel Accord was 26 pages long, whereas the final draft of Basel II exceeded 500 pages. The original timetable called for the completion of the final round of consultation by the end of 2001, with the new rules taking effect by 2004.

However, criticism from banks, trade associations, and national regulators led to several postponements. The final draft of Basel II was not published until June 2004, and the "Big Six" internationally active Canadian banks began to implement the new accord at the end of 2007, whereas European banks began to implement it at the start of 2008. U.S. banks submitted plans for compliance with Basel II in 2008, but full implementation did not occur until 2009. Only the dozen or so largest U.S. banks are subject to Basel II; all others will be allowed to use a simplified version of the standards it imposes.

The global financial crisis, however, revealed many limitations of the new accord. First, Basel II did not require banks to have sufficient capital to weather the kind of financial disruption that occurred during this period. Second, risk weights in the standardized approach are heavily reliant on credit ratings, which proved to be extremely unreliable in the run-up to the financial crisis. Third, Basel II is very procyclical—that is, it demands that banks hold less capital when times are good but more when times are bad, thereby exacerbating credit cycles. Because the probability of default and expected losses on different classes of assets rises during bad times, Basel II may require more capital at exactly the time when capital is most short. This has been a particularly serious concern in the aftermath of the global financial crisis. As a result of this crisis, banks' capital balances eroded, leading to a cutback on lending that was a big drag on the economy. Basel II made this cutback in lending even worse, doing yet more harm to the economy. Fourth, Basel II did not focus sufficiently on the dangers of a possible drying up of liquidity, a problem that brought down financial institutions during the financial crisis.

As a result of these limitations, in 2010 the Basel Committee developed a new accord, Basel III. It beefs up capital standards not only by raising them substantially but also by improving the quality of the capital, makes capital standards less procyclical by raising capital requirements in good times and lowering them in bad, makes new rules on the use of credit ratings, and requires financial institutions to have more stable funding so that they are better able to withstand liquidity shocks. Measures to achieve these objectives are highly controversial because of concerns that tightening up capital standards might cause banks to restrict their lending, which would make it harder for economies throughout the world to recover from the recent deep recession. Basel III was originally scheduled to be implemented by 2015, but its implementation has been extended repeatedly and now is slated for completion in January 2022. However, whether Basel III will be fully in place by that date, and whether it will be successful in restraining risk taking, is highly uncertain.

be granted if existing banks in a community would be hurt by its presence. Today this anticompetitive stance (justified by the desire to prevent bank failures of existing banks) is no longer as strong in the chartering agencies.

Once a bank has been chartered, it is required to file periodic (usually quarterly) *call reports* that reveal the bank's assets and liabilities, income and dividends, ownership, foreign exchange operations, and other details. The bank is also subject to examination by the bank regulatory agencies to ascertain its financial condition at least once a year. To avoid duplication of effort, the three federal agencies work together and usually accept each other's examinations. This means that, typically, chartered banks are examined by the Office of the Superintendent of Financial Institutions (OSFI), the Canada Deposit Insurance Corporation, and the Bank of Canada.

Regular on-site examinations, which allow regulators to monitor whether an institution is complying with capital requirements and restrictions on asset holdings, also function to limit moral hazard. Bank examiners give banks a *CAMELS rating*. The acronym is based on the six areas assessed: capital adequacy, asset quality, management, earnings, liquidity, and sensitivity to market risk. With this information about a bank's activities, regulators can enforce regulations by taking such formal actions as issuing *cease and desist orders* to alter the bank's behaviour or even closing a bank if its CAMELS rating is sufficiently low. Actions taken to reduce moral hazard by restricting banks from taking on too much risk help reduce the adverse selection problem further, because with less opportunity for risk taking, risk-loving entrepreneurs will be less likely to be attracted to the banking industry.

Note that the methods used by regulators to cope with adverse selection and moral hazard have their counterparts in private financial markets (see Chapters 8 and 12). Chartering is similar to the screening of potential borrowers; regulations that restrict risky asset holdings are similar to restrictive covenants that prevent borrowing firms from engaging in risky investment activities; capital requirements act like restrictive covenants that require minimum amounts of net worth for borrowing firms; and regular examinations are similar to the monitoring of borrowers by lending institutions.

Bank examinations are conducted by bank examiners, who study a bank's books to see whether it is complying with the rules and regulations that apply to its holdings of assets. If a bank is holding securities or loans that are too risky, the bank examiner can force the bank to get rid of them. If a bank examiner decides that a loan is unlikely to be repaid, the examiner can force the bank to declare the loan worthless (to write off the loan, which reduces the bank's capital). If, after examining the bank, the examiner feels that it does not have sufficient capital or has engaged in dishonest practices, the bank can be declared a "problem bank" and will be subject to more frequent examinations.

Assessment of Risk Management

Traditionally, on-site examinations have focused primarily on assessment of the quality of a financial institution's balance sheet at a point in time and whether it complies with capital requirements and restrictions on asset holdings. Although the traditional focus is important for reducing excessive risk taking by financial institutions, it is no longer thought to be adequate in today's world, in which financial innovation has produced new markets and instruments that make it easy for financial institutions and their employees to place huge bets easily and quickly. In this new financial environment, a financial institution that is healthy at a particular point in time can be driven into insolvency extremely rapidly from trading losses, as forcefully demonstrated by the failure of Barings in 1995 (to be discussed in Chapter 12). Thus an examination that

focuses only on a financial institution's position at a point in time may not be effective in indicating whether it will, in fact, take on excessive risk in the near future.

This change in the financial environment resulted in a major shift in thinking about the prudential supervisory process throughout the world. Bank examiners, for example, are now placing far greater emphasis on evaluating the soundness of a bank's management processes with regard to controlling risk. This shift in thinking is now reflected in a new focus on risk management in the guidelines to examiners on trading and derivatives activities. Now, bank examiners give a separate risk-management rating, from 1 to 5, that feeds into the overall management rating as part of the CAMELS system. Four elements of sound risk management are assessed to arrive at the risk-management rating: (1) the quality of oversight provided by the board of directors and senior management, (2) the adequacy of policies and limits for all activities that present significant risks, (3) the quality of the risk measurement and monitoring systems, and (4) the adequacy of internal controls to prevent fraud or unauthorized activities on the part of employees.

This shift toward focusing on management processes is also reflected in recent guidelines adopted by the Canadian bank regulatory authorities to deal with interest-rate risk. These guidelines require the bank's board of directors to establish interest-rate risk limits, appoint officials of the bank to manage this risk, and monitor the bank's risk exposure. The guidelines also require that senior management of a bank develop formal risk-management policies and procedures to ensure that the board of directors' risk limits are not violated and to implement internal controls to monitor interest-rate risk and compliance with the board's directives. Particularly important is the implementation of **stress tests**, which calculate losses under fictional dire scenarios and the need for more capital, and **value-at-risk (VaR)** calculations, which measure the size of the loss on a trading portfolio—say, over a two-week period—that might happen 1% of the time. In addition to these guidelines, bank examiners will continue to consider interest-rate risk in deciding the bank's capital requirements.

Disclosure Requirements

The free-rider problem described in Chapter 8 indicates that individual depositors and creditors will not have enough incentive to produce private information about the quality of a financial institution's assets. To ensure that better information is available in the marketplace, regulators can require that financial institutions adhere to certain standard accounting principles and disclose a wide range of information that helps the market assess the quality of an institution's portfolio and the amount of its exposure to risk. More public information about the risks incurred by financial institutions and the quality of their portfolios can better enable stockholders, creditors, and depositors to evaluate and monitor financial institutions and so act as a deterrent to excessive risk taking.

Disclosure requirements are a key element of financial regulation. Basel II puts a particular emphasis on disclosure requirements, with one of its three pillars focusing on increasing market discipline by mandating increased disclosure by banking institutions of their credit exposure, amount of reserves, and capital. Provincial securities commissions, the most significant being the Ontario Securities Commission (OSC), also impose disclosure requirements on any corporation, including financial institutions, that issues publicly traded securities. In addition, the OSC has required financial institutions to provide additional disclosure regarding their off-balance-sheet positions and more information about how they value their portfolios.

Regulation to increase disclosure is needed to limit incentives to take on excessive risk and to upgrade the quality of information in the marketplace so that investors

can make informed decisions, thereby improving the ability of financial markets to allocate capital to its most productive uses. The efficiency of markets is assisted by the OSC's disclosure requirements mentioned previously as well as its regulation of brokerage firms, mutual funds, exchanges, and credit-rating agencies to ensure that they produce reliable information and protect investors. Bill 198, enacted by the Ontario government in December 2002 in response to the **Sarbanes-Oxley Act** in the United States, took disclosure of information even further by increasing the incentives to produce accurate audits of corporate income statements and balance sheets and putting in place regulations to limit conflicts of interest in the financial services industry.

Consumer Protection

The existence of asymmetric information also suggests that consumers may not have enough information to protect themselves fully in financial dealing. Consumer protection regulation has taken several forms. First is "truth in lending," which requires all lenders, not just banks, to provide information to consumers about the cost of borrowing, including the disclosure of a standardized interest rate (called the *annual percentage rate* or *APR*) and the total finance charges on the loan. Legislation also requires creditors, especially credit card issuers, to provide information on the method of assessing finance charges and requires that billing complaints be handled quickly.

The global financial crisis illustrated the need for greater consumer protection, because so many borrowers took out loans with terms they did not understand and that were well beyond their means to repay. Most infamous were so-called NINJA loans in the United States, which were issued to borrowers with No Income, No Job, and No Assets. The result was millions of foreclosures, with many people losing their homes. Because weak consumer protection regulation played a prominent role in this crisis, there have been increasing demands to beef up this regulation, not only in Canada and the United States but in other countries as well.

Restrictions on Competition

Increased competition can also increase moral hazard incentives for financial institutions to take on more risk. Declining profitability resulting from increased competition could tip the incentives of financial institutions toward assuming greater risk in an effort to maintain former profit levels. Thus, governments in many countries have instituted regulations to protect financial institutions from competition. These regulations took different forms in the past, one being preventing nonbank institutions from competing with banks by engaging in banking business.

Although restrictions on competition propped up the health of banks, there were some serious disadvantages: they led to higher charges to consumers and decreased the efficiency of banking institutions, which now did not have to compete as vigorously. Thus, although the existence of asymmetric information provided a rationale for anticompetitive regulations, it did not mean that the regulations would be beneficial to consumers. Indeed, in recent years, the impulse of governments in industrialized countries to restrict competition has been waning.

Summary

Asymmetric information analysis explains what types of financial regulations are needed to reduce moral hazard and adverse selection problems in the financial system. However, understanding the theory behind regulation does not mean that regulation and supervision

of the financial system are easy in practice. Getting regulators and supervisors to do their job properly is difficult for several reasons. First, as we will see in our discussion of financial regulation in Chapter 10, financial institutions, in their search for profits, have strong incentives to avoid existing regulations by loophole mining. Thus, regulation applies to a moving target. Regulators are continually playing cat and mouse with financial institutions—financial institutions think up clever ways to skirt regulations, which leads regulators to modify their regulatory activities. Regulators continually face new challenges in a dynamically changing financial system, and unless they can respond rapidly, they may not be able to prevent financial institutions from taking on excessive risk. This problem can be exacerbated if regulators and supervisors do not have the resources or expertise necessary to keep up with clever people seeking to circumvent the existing regulations.

Financial regulation and supervision are difficult for two other reasons. In the regulation and supervision game, the devil is in the details. Subtle differences in the details may have unintended consequences; unless regulators get the regulation and supervision just right, they may be unable to prevent excessive risk taking. In addition, regulated firms may lobby politicians to lean on regulators and supervisors to go easy on them or to change regulations altogether.

Global International Financial Regulation

Because asymmetric information problems in the banking industry are a fact of life throughout the world, financial regulation in other countries is similar to that in Canada. Financial institutions are chartered and supervised by government regulators, just as they are in Canada. Disclosure requirements for financial institutions and corporations issuing securities are similar in other developed countries. Deposit insurance is also a feature of the regulatory systems of most other countries, although its coverage is different across countries. In the United States, for example, the FDIC guarantees that depositors will be paid off in full on the first $250 000 they have deposited in a bank if the bank fails. We have also seen that capital requirements are in the process of being standardized across countries in compliance with agreements like the Basel Accord.

Particular problems in financial regulation occur when financial institutions operate in many countries and thus can shift their business readily from one country to another. Financial regulators closely examine the domestic operations of financial institutions in their own country, but they often do not have the knowledge or ability to keep a close watch on operations in other countries, operations run either by domestic institutions' foreign affiliates or by foreign institutions with domestic branches. In addition, when a financial institution operates in many countries, it is not always clear which national regulatory authority should have primary responsibility for keeping the institution from engaging in overly risky activities.

The difficulties inherent in international financial regulation were highlighted by the collapse of the Bank of Credit and Commerce International (BCCI). BCCI, although it operated in more than 70 countries, including the United States and the United Kingdom, was supervised by Luxembourg, a tiny country unlikely to be up to the task. When massive fraud was discovered, the Bank of England closed BCCI down, but not before depositors and stockholders were exposed to huge losses. Cooperation among regulators in different countries and standardization of regulatory requirements provide potential solutions to the problems of international financial regulation.

The world has moved in this direction through agreements like the Basel Accord and oversight procedures announced by the Basel Committee in July 1992, which require a bank's worldwide operations to be under the scrutiny of a single home-country regulator with enhanced powers to acquire information on the bank's activities. The Basel Committee also ruled that regulators in other countries can restrict the operations of a foreign bank if they believe it lacks effective oversight. Whether agreements of this type will solve the problem of international financial regulation in the future is an open question.

For all these reasons, there is no guarantee that regulators and supervisors will be successful in promoting a healthy financial system. These same problems bedevil financial regulators in other countries as well, as the Global box "International Financial Regulation" indicates. Indeed, financial regulation and supervision have not always worked well, leading to banking crises in Canada and throughout the world.[1]

Because so many laws regulating the financial system have been passed in Canada, it is hard to keep track of them all. As a study aid, Table 9-1 lists the major financial legislation in the twentieth century and its key provisions.

TABLE 9-1 Major Financial Legislation in Canada

Bank of Canada Act (1934)

- Created the Bank of Canada following the recommendations of the Macmillan Commission of 1933

Bank Act of 1936

- Prohibited banks from issuing banknotes
- Imposed reserve requirements on depository institutions to be held with the Bank of Canada

Bank Act of 1954

- Increased reserve requirements on depository institutions
- Allowed chartered banks to offer mortgages issued under the National Housing Act

Canada Deposit Insurance Corporation Act (1967)

- Created the CDIC to insure deposits with all federally chartered banks and near banks

Bank Act of 1967

- Removed the 6% loan interest-rate ceiling
- Restricted foreign competition
- Imposed secondary reserve requirements on depository institutions
- Put chartered banks and near banks on equal footing regarding mortgage lending

Bank Act of 1981

- Created the Canadian Payments Association to operate the national payments system and plan its development
- Lowered reserve requirements on Canadian-dollar deposits
- Increased competition by introducing a less complicated procedure for obtaining a licence to operate as a bank
- Allowed foreign banks to establish subsidiaries in Canada, subject to reciprocal treatment of Canadian banks
- Extended banks' business powers to include financial leasing, factoring, and data processing
- Redesigned corporate clauses to ensure consistency between the Bank Act and the Canada Business Corporations Act
- Provided a simpler incorporation method for new banks (letters patent) while retaining incorporation through a special Act of Parliament

Office of the Superintendent of Financial Institutions Act (1987)

- Created the OSFI to succeed two separate federal regulatory bodies (the Department of Insurance and the Inspector General of Banks) in the supervision of financial institutions

Continued

[1]Canada suffered a major banking crisis in the 1980s, a crisis that hit the banking industry particularly hard. This crisis is discussed in the first online appendix to this chapter, "The 1980s Canadian Banking Crisis." A second online appendix to this chapter, "Banking Crises Throughout the World," discusses banking crises in other countries.

TABLE 9-1 *(continued)*

Financial Institutions and Deposit Insurance System Amendment Act (1987)

- Allowed chartered banks to own investment banking subsidiaries, thereby initiating the merging of the four pillars

Savings and Credit Union Act of the Province of Québec (1988)

- Set rules for credit unions, federations, and confederations, with specific reference to the *Mouvement Desjardins* (Desjardins Group)

Bank Act of 1992

- Comprehensive banking law
- Allowed chartered banks to own trust companies
- Allowed trust companies to make commercial loans
- Made provisions for the phasing out of reserve requirements
- Set rules regarding the supervisory role of the Bank of Canada and the OSFI with respect to chartered banks
- Reset the "sunset" clause from 10 years to 5 years to address the changing Canadian financial services marketplace

Cooperative Credit Associations Act (1992)

- Replaced the Cooperative Credit Association Act of 1952–1953 and set rules for federally chartered credit unions
- Followed the same format as the Bank Act

Insurance Companies Act (1992)

- Replaced the Canadian and British Insurance Companies Act and the Foreign Insurance Companies Act, both passed in 1932
- Set rules for life insurance companies and property and casualty (P&C) insurance companies
- Allowed insurance companies to own Schedule II chartered banks
- Followed the same format as the Bank Act

Trust and Loan Companies Act (1992)

- Replaced the Trust Companies Act and the Loan Companies Act, both passed in 1914
- Set rules for federally incorporated Trust and Mortgage Loan Companies (TMLs) and provincially incorporated TMLs reporting to the OSFI
- Required large, formerly closely held TMLs to become 35% widely held
- Followed the same format as the Bank Act

Bank Act of 1997

- Yielded minor changes because the government was waiting for the recommendations of the MacKay Task Force

Bank Act Reform of 2001

- Set new ownership rules
- Established a process for reviewing mergers involving large banks
- Allowed bank financial groups to organize under a holding company structure
- Allowed greater flexibility for bank involvement in the information technology area
- Allowed non-deposit-taking financial institutions access to the payments and clearance systems

Financial System Review Act (FSRA) of March 29, 2012

- Amendments to the Bank Act and other associated statutes

Amendments to the CDIC Act (in the aftermath of the global financial crisis of 2007–2009)

- In 2011, the CDIC became Canada's resolution authority
- In 2016, the "bail-in power" was added to the tools of the CDIC
- In 2020, expanded coverage of eligible deposits was introduced

9.3 CDIC DEPOSIT INSURANCE COVERAGE

LO 9.3 Examine how the Canada Deposit Insurance Corporation (CDIC) provides deposit insurance in Canadian banks and savings institutions.

The CDIC insures each depositor at member institutions up to a loss of $100 000 per account. All federally incorporated financial institutions and all provincially incorporated trust and mortgage loan companies are members of the CDIC. Insurance companies, credit unions, *caisses populaires*, and investment dealers are not eligible for CDIC membership; the Québec Deposit Insurance Board (QDIB) insures provincially incorporated financial institutions in Québec, and the other provinces have deposit insurance corporations that insure the deposits of credit unions in their jurisdiction, on terms similar to the CDIC's.

The CDIC manages about $1 trillion in credit risk in the form of insured deposits at 85 CDIC-member institutions—more than two-thirds of that is from the deposits of domestic systemically important banks (DSIBs) and the balance from non-DSIBs. Not all deposits and investments offered by CDIC-member institutions are insured by the CDIC. Insurable deposits include savings and chequing accounts, GICs and other term deposits, and foreign currency deposits. The CDIC does not insure mutual funds, stocks and bonds, Exchange Traded Funds (ETFs), and cryptocurrencies.

The primary rationale for deposit insurance is to protect depositors from bank insolvency and thus ensure financial stability. Deposit insurance could also promote competition among financial institutions by removing barriers to entry for new deposit-taking institutions. In the absence of deposit insurance, it is difficult for new banks to attract deposits. Most depositors, for example, are not capable of making appropriate risk calculations to assess the risk of a new bank. Those depositors would tend to place their deposits in banks that are considered too big to fail, thereby producing significant barriers to entry and unfair disadvantages for small new entrants. By insuring deposits at all deposit-taking financial institutions, the CDIC effectively removes barriers to entry for new deposit takers.

Differential Premiums

Until May 1999, CDIC premium revenue was not tied to the risk profile of financial institutions; the premium rate was the same for all deposit-taking institutions, irrespective of their risk profile. For example, in the 1998–1999 fiscal year, the flat-rate insurance premium was one-sixth of 1%, or 0.1667%, meaning that each deposit-taking financial institution paid an insurance premium of close to 17 cents per $100. This was one of the reasons that the Big Six, represented by the Canadian Bankers Association, vigorously opposed the establishment of the CDIC in 1967; it was argued that deposit insurance would be a subsidy to small banks paid by the big banks.

Over the years, the Canadian Bankers Association strongly promoted the reform of the Canadian deposit insurance system. As a result, the CDIC developed the Differential Premiums By-law, which came into effect for the premium year beginning May 1, 1999. The bylaw undergoes regular reviews and has been amended a number of times. The important feature of this legislation is its implicit, prompt corrective action provisions, which require the CDIC to intervene earlier and more vigorously when a bank gets into trouble. CDIC-member institutions are classified into one of four premium groups based on their risk profile. An institution's risk profile is determined using a variety of quantitative and qualitative criteria, including capital adequacy, profitability,

TABLE 9-2 Premium Structure and Rates for CDIC-Member Institutions in 2017

Category	Premium Rate (Basis Points % of Insured Deposits)
1	7.5
2	15.0
3	30.0
4	33.33

Source: CDIC website: www.cdic.ca

asset concentration, income volatility, regulatory ratings, and adherence to CDIC's Standards of Sound Business and Financial Practices, with capital adequacy dominating the criteria, accounting for 20% of the score.

Under this system, the premium rates for CDIC-member institutions are those shown in Table 9-2; they vary from 6.5 cents to 33.33 cents per $100. Category 1, classified as "best," is well-capitalized banks that significantly exceed minimum requirements. On the other hand, banks in Category 4, classified as "worst," are significantly (and perhaps critically) undercapitalized, and the insurance premium that they pay is over 33 basis points. In addition, for Category 4 banks, the CDIC is required to take prompt corrective actions such as requiring them to submit a capital restoration plan, restrict their asset growth, and seek regulatory approval to open new branches or develop new lines of business. Today, over 90% of CDIC-member institutions are classified in Categories 1 or 2, but, as in other countries, the premium category and related supervisory information applicable to individual CDIC members are confidential.

Opting Out

Another interesting development is the Opting-Out By-law that came into effect on October 15, 1999. This legislation permits Schedule III banks that primarily accept wholesale deposits (defined as $150 000 or more) to opt out of CDIC membership and therefore to operate without deposit insurance. This legislation, however, includes provisions to protect depositors who hold deposits eligible for CDIC protection. In particular, it requires an opted-out bank to inform all depositors, by posting notices in its branches, that their deposits will not be protected by the CDIC and not to charge any early withdrawal penalties for depositors who choose to withdraw.

Probably the most important feature of the opting-out legislation is its minimization of CDIC exposure to uninsured deposits. This represents a significant departure from past practices, when the CDIC showed generosity to uninsured depositors. For example, in the Canadian Commercial and Northland Bank failures of the mid-1980s, the CDIC paid 100 cents on the dollar to all depositors, both insured and uninsured. By compensating only the insured depositors rather than all depositors, the opting-out legislation increases the incentives of uninsured depositors to monitor the risk-taking activities of banks, thereby reducing moral hazard risk.

Resolution Authority and the Bail-in Power

Since it was created in 1967, the CDIC has dealt with 43 failures of member institutions. However, it has not experienced a failure since 1996. To address the factors that contributed to the global financial crisis of 2007–2009, the federal government amended the CDIC Act to designate the CDIC as resolution authority for all its member institutions—banks, federally regulated credit unions as well as loan and trust companies and associations governed by the Cooperative Credit Associations Act that take deposits. The CDIC is responsible for handling the failure of its member institutions with the objective of protecting eligible deposits, maintaining the flow of critical financial services in the economy, and minimizing the risk to taxpayers.

As Canada's resolution authority, the CDIC also adopted the bail-in regime, consistent with international standards set out by the Financial Stability Board (an international body that promotes global financial stability). Bail-in is an important tool that allows the CDIC to ensure failing institutions remain open while losses are covered by shareholders and certain investors, in contrast to a bail-out that relies on taxpayer-funded capital injections to support failing institutions.

SUMMARY

1. The concepts of asymmetric information, adverse selection, and moral hazard help explain why governments pursue financial regulation.
2. There are eight basic types of financial regulation aimed at lessening asymmetric information problems and excessive risk taking in the financial system: (1) restrictions on asset holdings, (2) capital requirements, (3) prompt corrective action, (4) chartering and examination, (5) assessment of risk management, (6) disclosure requirements, (7) consumer protection, and (8) restrictions on competition.

MyLab Economics Make the grade with MyLab Economics: Select end-of-chapter questions can be found on MyLab Economics. You can practise them as often as you want, and most feature step-by-step guided instructions to help you find the right answer.

QUESTIONS

1. Why are deposit insurance and other types of government safety nets important to the health of the economy?
2. If casualty insurance companies provided fire insurance without any restrictions, what kind of adverse selection and moral hazard problems might result?
3. Do you think that eliminating or limiting the amount of deposit insurance would be a good idea? Explain your answer.
4. How could higher deposit insurance premiums for banks with riskier assets benefit the economy?
5. What are the costs and benefits of a too-big-to-fail policy?
6. Suppose that you have $300 000 in deposits at a bank. After careful consideration, the FDIC decides that this bank is now insolvent. Which method would you like to see the FDIC apply? What if your deposit were $200 000?

7. Would you recommend the adoption of a system of deposit insurance, like the CDIC in Canada, in a country with weak institutions, prevalent corruption, and ineffective regulation of the financial sector?
8. At the height of the global financial crisis in October 2008, the U.S. Treasury forced nine of the largest U.S. banks to accept capital injections in exchange for non-voting ownership stock, even though some of the banks did not need the capital and did not want to participate. What could the rationale have been for doing this?
9. What special problem do off-balance-sheet activities present to bank regulators, and what have they done about it?
10. What are some of the limitations to the Basel and Basel II Accords? How does the Basel III Accord attempt to address these limitations?
11. How does bank chartering reduce the adverse selection problems? Does it always work?
12. Why has the trend in bank supervision moved away from a focus on capital requirements to a focus on risk management?
13. Suppose that after a few mergers and acquisitions, only one bank holds 70% of all deposits in Canada. Would you say that this bank would be considered too big to fail? What does this tell you about the ongoing process of financial consolidation and the government safety net?
14. Suppose Universal Bank holds $100 million in assets, which are composed of the following:

Desired reserves:	$10 million
Excess reserves:	$ 5 million
Mortgage loans:	$20 million
Corporate bonds:	$15 million
Stocks:	$25 million
Commodities:	$25 million

Do you think it is a good idea for Universal Bank to hold stocks, corporate bonds, and commodities as assets? Why or why not?

15. Why might more competition in financial markets be a bad idea? Would restrictions on competition be a better idea? Why or why not?
16. In what way might consumer protection regulations negatively affect a financial intermediary's profits? Can you think of a positive effect of such regulations on profits?

APPLIED PROBLEMS

17. Consider a failing bank. A deposit of $350 000 is worth how much to the depositor if the CDIC uses the *payoff* method given the typical recovery rate? How much is the same deposit worth to the depositor if the *purchase-and-assumption* method is used? Which is more costly to taxpayers?
18. Consider a bank with the balance sheet below:
Under the original Basel Accord, calculate the bank's risk-weighted assets.
19. Oldhat Financial starts its first day of operations with $9 million in capital. A total of $130 million in chequable deposits is received. The bank makes a $25 million commercial loan and another $50 million in mortgages with the following terms: 200 standard, 30-year, fixed-rate mortgages with a nominal annual rate of 5.25%, each for $250 000.

Assume that desired reserves are 8%.

a. What does the bank balance sheet look like?

b. How well capitalized is the bank?

c. Calculate the risk-weighted assets and risk-weighted leverage ratio after Oldhat's first day.

20. Early the next day, Oldhat Financial invests $50 million of its excess reserves in commercial loans. Later

Assets		Liabilities	
Desired reserves	$17 million	Chequable deposits	$205 million
Excess reserves	$ 3 million	Bank capital	$ 10 million
Municipal bonds	$65 million		
Residential mortgages	$70 million		
Commercial loans	$60 million		

that day, terrible news hits the mortgage markets, and mortgage rates jump to 13%, implying a present value of Oldhat's current mortgage holdings of $124 798 per mortgage. Bank regulators force Oldhat to sell its mortgages to recognize the fair market value. What does Oldhat's balance sheet look like? How do these events affect its capital position?

21. To avoid insolvency, regulators decide to provide Oldhat Financial with $25 million in bank capital. However, the bad news about the mortgages is featured in the local newspaper, causing a bank run. As a result, $30 million in deposits is withdrawn. Show the effects of the capital injection and the bank run on the balance sheet. Was the capital injection enough to stabilize the bank? If the bank regulators decide that the bank needs a leverage ratio of 10% to prevent further runs on the bank, how much of an additional capital injection is required to reach a 10% leverage ratio?

DATA ANALYSIS PROBLEMS

1. Go to the CANSIM database and find monthly data on total chartered bank assets (series V36883), total liabilities and shareholders' equity (series V36936), and shareholders' equity (series V36960 + V36961 + V36962 + V36964 + V41598372 + V53843372). Download the data from January 2011 through the most recent month available into a spreadsheet and plot the series. For each monthly observation, calculate the bank leverage ratio as the ratio of bank capital to total assets. Create a line graph of the leverage ratio over time. All else being equal, what can you conclude about leverage and moral hazard in Canadian chartered banks over time?

2. **Real-Time Data Analysis** Go to the St. Louis Federal Reserve FRED database and find data on the number of commercial banks in the United States in each of the following categories: average assets less than $100 million (US100NUM), average assets between $100 million and $300 million (US13NUM), average assets between $300 million and $1 billion (US31NUM), average assets between $1 billion and $15 billion (US115NUM), and average assets greater than $15 billion (USG15NUM). Download the data into a spreadsheet. Calculate the percentage of banks in the smallest (less than $100 million) and largest (greater than $15 billion) categories as a percentage of the total number of banks for the most recent quarter of data available and for 1990:Q1. What happened to the proportion of very large banks? What has happened to the proportion of very small banks? What does this say about the too-big-to-fail problem and moral hazard?

10 Banking Industry: Structure and Competition

Learning Objectives

10.1 Recognize the key features of the Canadian banking system and the historical context of the implementation of these features.

10.2 Explain how financial innovation led to the growth of the shadow banking system.

10.3 Identify the key structural changes in the chartered banking industry.

10.4 Identify the key differences between the Canadian and U.S. banking industries.

10.5 Summarize the factors that led to consolidation in the chartered banking industry.

10.6 Summarize the distinctions between chartered banks and trust and loan companies and credit unions and *caisses populaires*.

10.7 Identify the reasons for Canadian banks to operate in foreign countries and for foreign banks to operate in Canada.

Preview

The operations of individual banks (how they acquire, use, and manage funds to make a profit) are roughly similar throughout the world. In all countries, banks are financial intermediaries in the business of earning profits. When you consider the structure and operation of the banking industry as a whole, however, Canada is in a class by itself, with a comparatively concentrated banking industry and small financial institutions. Unlike the United States, where there are on the order of 4500 commercial banks, 700 savings institutions (savings and loan associations and mutual savings banks), and 5000 credit unions, Canada has six large commercial banks, commonly called chartered banks, that typically dominate the banking industry.[1]

Is fewer better? Does it mean that the Canadian banking system is more stable and competitive and therefore more economically efficient and sound than banking systems in other countries? What in the Canadian economic and political system explains this small number of banking institutions? In this chapter we try to answer these questions by examining the historical trends in the banking industry and its overall structure.

We start by examining the historical development of the banking system and how financial innovation has increased the competitive environment for the banking industry and is causing fundamental changes in it. We then go on to look at the chartered banking industry in detail and then discuss the near-banking industry, which includes trust and loan companies and credit unions and *caisses populaires*. We spend more time on chartered banks because they are by far the largest depository institutions, accounting for over two-thirds of the deposits in the banking system. In addition to looking at our domestic banking system, we examine the forces behind the growth of international banking to see how it has affected us in Canada.

10.1 HISTORICAL DEVELOPMENT OF THE CANADIAN BANKING SYSTEM

LO 10.1 Recognize the key features of the banking system and the historical context of the implementation of these features.

The modern Canadian banking industry began with the creation of the Bank of Montreal in 1817 by nine merchants in Montreal. Initially, the Bank of Montreal was without statutory authority, but a charter was approved by the legislature of Lower Canada

[1]Commercial banks in Canada are called chartered banks because they can be established only by charter granted either in a special act of the federal Parliament or by the federal minister of finance.

10.8 Assess the reasons for separating banking from other financial services through legislation.

and confirmed by royal assent in 1822. Meanwhile, other banks opened for business, and the Canadian banking industry was off and running; the Bank of New Brunswick received royal assent in 1820 and the Chartered Bank of Upper Canada in York (Toronto) in 1821. (As a study aid, Figure 10-1 provides a timeline of the most important dates in the history of Canadian commercial banking before World War II.)

All these banks were authorized to issue notes (redeemable in *specie*, essentially British or American gold or silver coins, on demand), receive deposits, and lend for commercial purposes only; no bank was allowed to lend funds on mortgages, land, or real property. There were, however, some differences between the charters of these banks. For example, the terms of the charter of the Bank of New Brunswick followed the banking tradition of New England. The charter of the Bank of Montreal almost duplicated the terms governing the First Bank of the United States (see the FYI box "The Dual Banking System in the United States"), which had elements of both a private and a **central bank**, a government institution that has responsibility for the amount of money and credit supplied in the economy as a whole. Also, the Bank of New Brunswick had to submit regular annual statements to the government and was not allowed to open **branches** (additional offices for the conduct of banking operations), whereas

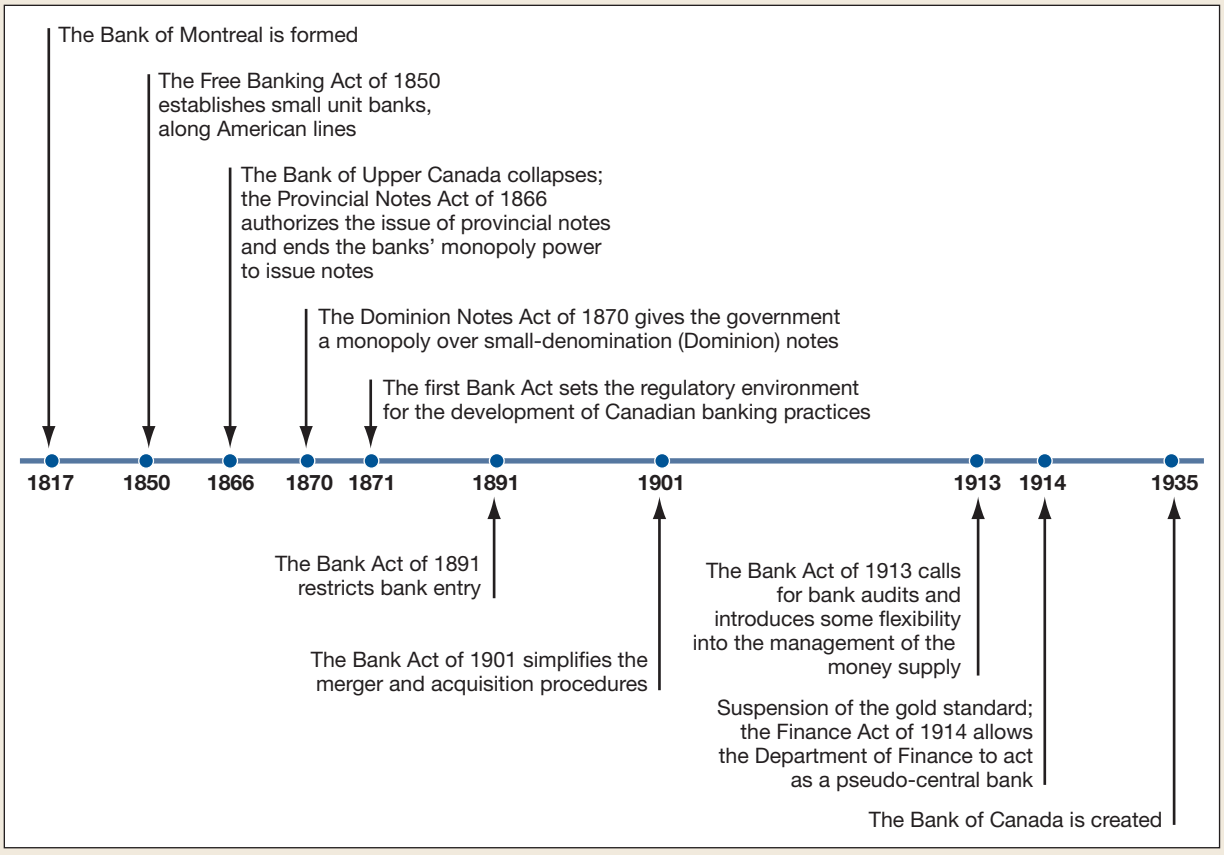

The Bank of Montreal is formed
The Free Banking Act of 1850 establishes small unit banks, along American lines
The Bank of Upper Canada collapses; the Provincial Notes Act of 1866 authorizes the issue of provincial notes and ends the banks' monopoly power to issue notes
The Dominion Notes Act of 1870 gives the government a monopoly over small-denomination (Dominion) notes
The first Bank Act sets the regulatory environment for the development of Canadian banking practices
1817
1850
1866
1870
1871
1891
1901
1913
1914
1935
The Bank Act of 1891 restricts bank entry
The Bank Act of 1901 simplifies the merger and acquisition procedures
The Bank Act of 1913 calls for bank audits and introduces some flexibility into the management of the money supply
Suspension of the gold standard; the Finance Act of 1914 allows the Department of Finance to act as a pseudo-central bank
The Bank of Canada is created

FIGURE 10-1 Timeline of the Early History of Commercial Banking in Canada

FYI The Dual Banking System in the United States

The modern commercial banking industry in the United States began when the Bank of North America was chartered in Philadelphia in 1782. A major controversy involving the U.S. banking industry in its early years was whether the federal government or the states should charter banks. The Federalists, particularly Alexander Hamilton, advocated greater centralized control of banking and federal chartering of banks. Their efforts led to the creation in 1791 of the First Bank of the United States.

Until 1863, all commercial banks in the United States were chartered by the banking commission of the state in which each operated. No national currency existed, and banks obtained funds primarily by issuing banknotes (currency circulated by the banks that could be redeemed for gold).

Because banking regulations were extremely lax in many states, banks regularly failed due to fraud or lack of sufficient bank capital; their banknotes became worthless. To eliminate the abuses of the state-chartered banks (called **state banks**), the National Bank Act of 1863 (and subsequent amendments to it) created a new banking system of federally chartered banks (called **national banks**). This legislation was originally intended to dry up sources of funds to state banks by imposing a prohibitive tax on their banknotes while leaving the banknotes of the federally chartered banks untaxed. The state banks cleverly escaped extinction by acquiring funds through deposits.

As a result, today the United States has a **dual banking system** in which banks chartered by the federal government and banks chartered by the states operate side by side.

the Bank of Montreal had to provide statements only on request and was allowed to open branches in any part of Upper or Lower Canada.

The Free Banking Experiment

Until 1850, no national currency existed in pre-Confederation Canada, and banks obtained funds primarily by issuing *banknotes* (currency circulated by the banks that could be redeemed for gold). No banks failed, but because banking regulations were extremely lax, banks regularly experienced substantial declines in bank capital due to business failures; their banknotes tended to become scarce. The government tried various schemes to guarantee the provision of stable money, including issuing its own notes, but it was significantly influenced by the concept of **free banking**, implemented in New York in 1837. This system, as the name suggests, permitted the organization of a bank by any group that met certain established criteria concerning the amount of equity capital and maintenance of reserves.

The Free Banking Act was passed in Canada in 1850, with the purpose of facilitating the entry of small unit banks along American lines. It allowed the establishment of a bank, without a legislative charter, by any group that met the lax requirements set out in the free banking legislation. Under this legislation, the minimum amount of net worth to organize a bank was $100 000, branching was not allowed, and although the banknotes of the free banks were untaxed, the amount of note issue was limited to the amount of government debt held by the banks. The move to free banking was a step in the right direction, but Canada's experience with free banking was a failure. It did not lead to the establishment of a large number of new banks; only five new banks were established, two of which soon failed, and the other three converted to legislative charters.

The restriction on branching and the issue of banknotes based on government debt, rather than on commercial loans, were blamed for the failure of Canada's free banking experiment. The most important factor, however, was the fact that the option of a legislative charter was still available, unlike the situation in the United States, where the provision of a legislative charter was abolished in those states where free banking was established. In Canada, free banking, with its restrictive provisions, particularly the restriction on branches and the less liberal provision for note issue, proved to be less profitable than banking under legislative charters.

In 1850, there were 15 chartered banks in Canada: 8 in Central Canada and 7 in what was to become Atlantic Canada. From 1850 until Confederation in 1867, and except for a short period after 1857, the Canadian provinces experienced an economic expansion, and 30 new banks were established. However, 11 of these failed or closed their doors for other reasons, leaving 34 chartered banks with a total of 127 branches at the end of 1867.

The Provincial Notes Act, 1866

In the years before Confederation, governments were anxious about the chartered banks' control of note issue. They believed that the best way to protect the public from some of the consequences of bank failures would be to separate the currency of the country from the banking interests. In 1860, Alexander Galt, finance minister of the Province of Canada, proposed the substitution of a government-issued paper currency for banknotes. His proposal, however, was defeated by his critics, especially the chartered banks, for obvious reasons: The substitution of interest-free government debt for interest-free bank debt would have directly reduced their profits.

In the midst of a minor financial crisis in 1866, with the collapse of the Bank of Upper Canada (Canada's first chartered bank failure), the proponents of government-issued paper money finally achieved their objective with the enactment of the Provincial Notes Act. The act authorized the issue of provincial notes, which, because of their legal reserve status, could be substituted for specie. With the cooperation of the Bank of Montreal, which had become the government's fiscal agent in 1864 by replacing the Bank of Upper Canada, the banks began to hold the new currency, thereby surrendering their power to issue notes.

The Dominion Notes Act, 1870

Canada was created by the Constitution Act (formerly the British North America Act) in 1867. The act granted the new federal government of Canada exclusive jurisdiction over all matters pertaining to currency and banking, and the first problem to be tackled was the issue of paper money. With the failure in 1867 of the Chartered Bank of Canada (the second chartered bank failure in Canadian history), the Dominion Notes Act was passed in 1870. The act confirmed the rights of banks to issue banknotes on their own credit but restricted them to large-denomination (over $5) notes, thereby giving the government a monopoly over small-denomination ($1 and $2) notes, the Dominion notes.

Although the Dominion Notes Act of 1870 did not set any reserve requirements, it required banks to hold at least half of their reserves in Dominion notes, thereby giving the government a share of the profits from the issuance of money, which is called **seigniorage**. The Dominion notes themselves were fractionally backed by gold, and in this sense the Dominion Notes Act of 1870 confirmed that Canada would operate under the **gold standard**, meaning that its currency was convertible directly into gold.

Canada operated under the gold standard, keeping its currency backed by and convertible into gold, until World War I. During the years 1870–1935, Dominion notes increased in importance, but they never accounted for a major fraction of currency in circulation. They were superseded, together with the banknotes, by Bank of Canada notes soon after the creation of the Bank of Canada (Canada's central bank) in 1935.

The First Bank Act, 1871

The first Bank Act came into effect in 1871. It was to be revised every 10 years, in light of experience and changing conditions; this "sunset" clause effectively ensured that governments over the years paid periodic attention to banking reform. The Bank Act set the regulatory environment for Canadian chartered banks and for the future development of Canadian banking practices.

The act continued the legislative chartering of banks, with each charter running for a 10-year period, then to be reviewed and renewed. New banks had to meet minimum capital requirements: $100 000 paid up before they opened for business against a total of $500 000. The banks' note issue continued to be restricted to large-denomination notes (over $5) and limited to the amount of their paid-up capital plus reserves. There were no reserve requirements, but one-third of a bank's cash reserves were required to be in the form of Dominion notes.

The act continued the prohibition against mortgage lending and real estate loans, but it reinforced the commercial nature of banking by allowing banks to make loans on the security of most kinds of merchandise. Also, for the greater security of the public, bank shareholders were liable for double the amount of their subscription. Finally, each bank was required to submit a detailed statement to the government on a monthly basis, but there was no provision for government inspection or audit.

The Bank Act, 1881–1913

A depression followed Confederation and lasted from 1873 to 1879. During the depression years, the banks were hard hit, and 13 bank failures (four in 1878, five in 1887, and another four in 1890) wiped out the savings of many noteholders. To prevent future losses from such failures, the early decennial revisions of the Bank Act, in 1881, 1891, 1901, and 1913 (postponed since 1911), were intended to provide better protection for the holders of banknotes, but the act was not substantially changed.

In particular, in the Bank Act revision of 1891, the capital requirement was increased to $250 000 paid up, thereby restricting entry into the industry. The proportion of Dominion notes in bank cash reserves was increased to 40%, and the notes of a failed bank were made a first charge against its assets in the event of liquidation. Moreover, in the Bank Act revision of 1891, to insure noteholders against loss, a Bank Circulation Redemption Fund was created, with each bank contributing an amount equal to 5% of its average note circulation.

In the years between the mid-1890s and the outbreak of World War I, the Canadian economy experienced a phenomenal economic expansion. While bank entry was restrained (due to the increase in capital requirements in the Bank Act of 1891), the Bank Act revision of 1901 simplified the merger and acquisition procedures by requiring only approval of Cabinet; previously, a special Act of Parliament was required for all mergers. As a result of these legislative changes, 13 mergers took place before the end of 1914, compared to only 6 in the previous 33-year period, and the number of banks declined from 41 in 1890 to 22 in 1914. Over the same period, the number of bank branches increased from 426 to over 3000.

Another important legislative change occurred in the 1913 revision of the Bank Act. The act called for a *bank audit*: an annual, independent verification of the financial statements of the banks, with the results distributed to the shareholders and the minister of finance. The objective was to limit the adverse selection and moral hazard problems that had increased over the years and been found to be the cause of a number of bank failures, particularly the failure of the Farmers Bank in 1910.

An additional noteworthy change was the excess circulation provision that introduced some flexibility in the management of the money supply. The economic expansion in the period after the mid-1890s caused banknote issues to reach the ceiling that the Bank Act of 1871 had fixed at the amount of paid-up capital plus reserves. The banks did not increase their capital (and thus their note-issuing capacity), producing a shortage of currency. In order to achieve expansion in the money supply with the growth of economic activity, the Bank Act of 1913 allowed for the issuing of banknotes in excess of a bank's paid-up capital plus reserves.

The Finance Act, 1914

At the end of July 1914, less than a year after the revision of the Bank Act in 1913, World War I looked more and more inescapable. Canada's established banking legislation appeared to be inadequate, and the immediate problem was to preserve the stability and liquidity of the financial system. Panic had taken hold, with depositors converting their money into gold for hoarding and the banks and the government becoming concerned about their ability to convert money into gold on demand, since their gold reserves were a small fraction of their combined monetary liabilities. In light of these developments, on August 3, 1914, the government suspended the convertibility of Dominion notes and banknotes into gold, thereby ending the gold standard that had emerged over 40 years earlier in 1870. The gold standard was re-established in 1926 and suspended again in 1929, when the Great Depression hit the world.

A major legislative change, following the suspension of the gold standard, was the Finance Act of 1914. Patterned on the episode of 1907, during which banks could obtain cash reserves from the Department of Finance to prevent bank runs (which were triggered by bank failures in the United States), the Finance Act allowed the Department of Finance to act as a **lender of last resort**—that is, to provide Dominion notes to banks (on the pledge of approved securities) when no one else would, thereby preventing bank and financial panics. The Finance Act foreshadowed the increased flexibility in the management of the money supply that was provided by the Bank of Canada in 1935. We will examine in detail the economic forces that led to the creation of the Bank of Canada in Chapter 14.

10.2 FINANCIAL INNOVATION AND THE GROWTH OF THE "SHADOW BANKING SYSTEM"

LO 10.2 Explain how financial innovation led to the growth of the shadow banking system.

Although banking institutions are still the most important financial institutions in the Canadian economy, in recent years the traditional banking business of making loans that are funded by deposits has been in decline. Some of this business has been replaced

by the **shadow banking system**, in which bank lending has been replaced by lending via the securities markets, with the involvement of a number of different financial institutions.

To understand how the banking industry has evolved over time, we must first understand the process of financial innovation, which has transformed the entire financial system. Like other industries, the financial industry is in business to earn profits by selling its products. If a soap company perceives that a need exists in the marketplace for a laundry detergent with fabric softener, it develops a product to fit the need. Similarly, to maximize their profits, financial institutions develop new products to satisfy their own needs as well as those of their customers; in other words, innovation—which can be extremely beneficial to the economy—is driven by the desire to get (or stay) rich. This view of the innovation process leads to the following simple analysis: ***a change in the financial environment will stimulate a search by financial institutions for innovations that are likely to be profitable***.

Starting in the 1960s, individuals and financial institutions operating in financial markets were confronted with drastic changes in the economic environment: Inflation and interest rates climbed sharply and became harder to predict, a situation that changed demand conditions in financial markets. The rapid advances in computer technology changed supply conditions. In addition, financial regulations became more burdensome. Financial institutions found that many of the old ways of doing business were no longer profitable; the financial services and products they had been offering to the public were not selling. Many financial intermediaries found that they were no longer able to acquire funds with their traditional financial instruments, and without these funds they would soon be out of business. To survive in the new economic environment, financial institutions had to research and develop new products and services that would meet customer needs and prove profitable, a process referred to as **financial engineering**. In their case, necessity was the mother of innovation.

Our discussion of the factors that drive financial innovation suggests that there are three basic types of financial innovation: responses to changes in demand conditions, responses to changes in supply conditions, and avoidance of existing regulations. These three motivations often interact to produce particular financial innovations. Now that we have a framework for understanding why financial institutions produce innovations, let's look at examples of how financial institutions, in their search for profits, have produced financial innovations of the three basic types.

Responses to Changes in Demand Conditions: Interest-Rate Volatility

The most significant change in the economic environment that altered the demand for financial products in recent years has been the dramatic increase in the volatility of interest rates. In the 1950s, the interest rate on three-month Treasury bills fluctuated between 1.0% and 5.5%; in the 1970s, it fluctuated between 3% and 14%; in the 1980s, it ranged from 7% to more than 20%. Large fluctuations in interest rates lead to substantial capital gains or losses and greater uncertainty about returns on investments. Recall that the risk that is related to the uncertainty about interest-rate movements and returns is called *interest-rate risk*, and high volatility of interest rates, such as we saw in the 1970s and 1980s, leads to a higher level of interest-rate risk.

We would expect the increase in interest-rate risk to increase the demand for financial products and services that could reduce that risk. This change in the economic environment would thus stimulate a search for profitable innovations by financial

institutions that meet this new demand and would spur the creation of new financial instruments that help lower interest-rate risk. Two financial innovations that were developed in the 1970s confirm this prediction: adjustable-rate mortgages and financial derivatives.

Adjustable-Rate Mortgages Like other investors, financial institutions find that lending is more attractive if interest-rate risk is lower. They would not want to make a mortgage loan at a 10% interest rate and two months later find that they could obtain 12% in interest on the same mortgage. To reduce interest-rate risk, financial institutions began to issue adjustable-rate mortgages—that is, mortgage loans on which the interest rate changes when a market interest rate (usually the Treasury bill rate) changes. Initially, an adjustable-rate mortgage might have a 5% interest rate. In six months, this interest rate might increase or decrease by the amount of the increase or decrease in, say, the six-month Treasury bill rate, and the mortgage payment would then change. Because adjustable-rate mortgages allow mortgage-issuing institutions to earn higher interest rates on existing mortgages when market rates rise, profits remain high during these periods.

This attractive feature of adjustable-rate mortgages has encouraged mortgage-issuing institutions to issue adjustable-rate mortgages with lower initial interest rates than those on conventional fixed-rate mortgages, making them popular with many households. However, because the mortgage payment on a variable-rate mortgage can increase, many households continue to prefer fixed-rate mortgages. Hence, both types of mortgages are widespread.

Financial Derivatives Given the greater demand for the reduction of interest-rate risk, commodity exchanges such as the Chicago Board of Trade recognized that if they could develop a product that would help investors and financial institutions protect themselves from, or to **hedge**, interest-rate risk, then they could make profits by selling this new instrument. **Futures contracts**, in which the seller agrees to provide a certain standardized commodity to the buyer on a specific future date at an agreed-on price, had been around for a long time. Officials at the Chicago Board of Trade realized that if they created futures contracts in financial instruments, which are called **financial derivatives** because their payoffs are linked to (i.e., derived from) previously issued securities, they could be used to hedge risk. Thus, in 1975, financial derivatives were born.

Responses to Changes in Supply Conditions: Information Technology

The most important source of the changes in supply conditions that stimulate financial innovation has been the improvement in computer and telecommunications technology. This technology, called *information technology*, has had two effects. First, it has lowered the cost of processing financial transactions, making it profitable for financial institutions to create new financial products and services for the public. Second, it has made it easier for investors to acquire information, thereby making it easier for firms to issue securities. The rapid developments in information technology have resulted in many new financial products and services that we examine here.

Bank Credit and Debit Cards Credit cards have been around since well before World War II. Many individual stores (such as, for example, Home Depot and the Bay)

institutionalized charge accounts by providing customers with credit cards that allowed them to make purchases at these stores without cash. Nationwide credit cards were not established until after World War II, when Diners Club developed a card to be used in restaurants. Similar credit card programs were started by American Express and Carte Blanche, but because of the high cost of operating these programs, cards were issued only to selected persons and businesses that could afford expensive purchases.

A firm issuing credit cards earns income from loans it makes to credit card holders and from payments made by stores on credit card purchases (a percentage of the purchase price, say, 3%). A credit card program's costs arise from loan defaults, stolen cards, and the expense involved in processing credit card transactions.

Seeing the success of Diners Club, American Express, and Carte Blanche, bankers wanted to share in the profitable credit card business. Several chartered banks attempted to expand the credit card business to a wider market in the 1950s, but the cost per transaction of running these programs was so high that their early attempts failed.

In the late 1960s, improved computer technology, which lowered the transaction costs for providing credit card services, made it more likely that bank credit card programs would be profitable. The banks tried to enter this business again, and this time their efforts led to the creation of two successful bank credit card programs: Visa and MasterCard. These programs have become phenomenally successful; around 1.5 billion of their cards are in use around the world today. Indeed, bank credit cards have been so profitable that nonfinancial institutions such as General Motors and Walmart have also entered the credit card business. Consumers have benefited because credit cards are more widely accepted than cheques as a method of paying for purchases (particularly abroad), and they allow consumers to take out loans more easily.

The success of bank credit cards led these institutions to come up with another financial innovation, *debit cards*. Debit cards often look just like credit cards and can be used to make purchases in an identical fashion. However, in contrast to credit cards, which extend the purchaser a loan that does not have to be paid off immediately, a debit card purchase is immediately deducted from the card holder's bank account. Debit cards depend even more on low costs of processing transactions, because their profits are generated entirely from the fees paid by merchants on debit card purchases at their stores. Debit cards have grown extremely popular in recent years.

Electronic Banking The wonders of modern computer technology have also enabled banks to lower the cost of bank transactions by having the customer interact with an electronic banking (e-banking) facility rather than with a human being. One important form of an e-banking facility is the **automated teller machine (ATM)**, an electronic machine that allows customers to get cash, make deposits, transfer funds from one account to another, and check balances. ATM machines do not have to be paid overtime and never sleep and thus are available for use 24 hours a day. Not only does this result in cheaper transactions for the bank, but it also provides more convenience for the customer. Because of their low cost, ATMs can be put at locations other than a bank or its branches, further increasing customer convenience. The low cost of ATMs has encouraged banks to install them everywhere. Furthermore, it is now just as easy to get foreign currency from an ATM when you are travelling in Europe as it is to get cash from your local bank.

With the drop in the cost of telecommunications, banks have developed another financial innovation, *home banking*. It is now cost-effective for banks to set up an electronic banking facility in which the bank's customer is linked up with the bank's

computer and allowed to carry out transactions by using a smartphone, tablet, or personal computer. Now a bank's customers can conduct many of their bank transactions without ever leaving the comfort of home. The advantage for the customer is the convenience of home banking, while banks find that the cost of transactions is substantially less than the cost associated with bricks-and-mortar banking.

With the decline in the price of personal computers and their increasing presence in the home, we have seen a further innovation in the home banking area, the creation of a new type of banking institution, the **virtual bank**, a bank that has no physical location but rather exists only in cyberspace. In 1995, Security First Network Bank, based in Atlanta, Georgia, but now owned by Royal Bank of Canada, became the first virtual bank, offering an array of banking services exclusively on the Internet—accepting chequing account and savings deposits, selling certificates of deposit, issuing ATM cards, providing bill-paying facilities, and so on. The virtual bank thus takes home banking one step further, enabling the customer to have a full set of banking services at home, 24 hours a day. In 1996, Bank of America and Wells Fargo entered the virtual banking market, to be followed by many others, with Bank of America now being the largest Internet bank in the United States.

Junk Bonds Before the advent of computers and advanced telecommunications, it was difficult to acquire information about the financial situation of firms that might want to sell securities. Because of the difficulty in screening out bad from good credit risks, the only firms that were able to sell bonds were very well-established corporations that had high credit ratings.[2] Before the 1980s, then, only corporations that could issue bonds with ratings of BBB or above could raise funds by selling newly issued bonds. Some firms that had fallen on bad times, known as *fallen angels*, had previously issued long-term corporate bonds with ratings that had now fallen below BBB, bonds that were pejoratively dubbed "junk bonds."

With the improvement in information technology in the 1970s, it became easier for investors to acquire financial information about corporations, making it easier to screen out bad from good credit risks. With easier screening, investors were more willing to buy long-term debt securities from less-well-known corporations with lower credit ratings. With this change in supply conditions, we would expect that some smart individual would pioneer the concept of selling new public issues of junk bonds, not for fallen angels but for companies that had not yet achieved investment-grade status. This is exactly what Michael Milken of Drexel Burnham Lambert, an investment-banking firm, started to do in 1977. Junk bonds became an important factor in the corporate bond market. Although there was a sharp slowdown in activity in the junk bond market after Milken was indicted for securities law violations in 1989, the junk bond market heated up again in the 1990s and 2000s, with the amount outstanding now exceeding US$1 trillion.

Commercial Paper Market Recall that *commercial paper* is a short-term debt security issued by large banks and corporations. The commercial paper market underwent tremendous growth since the 1970s, with commercial paper becoming one of the most important money market instruments.

Improvements in information technology also help explain the rapid rise of the commercial paper market. We have seen that improvements in information technology made it easier for investors to screen out bad from good credit risks, thus making it easier for

[2]The discussion of adverse selection problems in Chapter 8 provides a more detailed analysis of why only well-established firms with high credit ratings were able to sell securities.

corporations to issue debt securities. Not only did this make it simpler for corporations to issue long-term debt securities, as in the junk bond market, but it also meant they could raise funds by issuing short-term debt securities, such as commercial paper, with greater ease. Many corporations that used to do their short-term borrowing from banks now frequently raise short-term funds in the commercial paper market instead.

The development of money market mutual funds has been another factor in the rapid growth of the commercial paper market. Because money market mutual funds need to hold liquid, high-quality, short-term assets such as commercial paper, the growth of assets in these funds has created a ready market in commercial paper. The growth of pension and other large funds that invest in commercial paper has also stimulated the growth of this market.

Securitization and the Shadow Banking System

One of the most important financial innovations of the past two decades, arising from improvements in information technology, is securitization. **Securitization** is the process of bundling small and otherwise illiquid financial assets (such as residential mortgages, auto loans, and credit card receivables), which have typically been the bread and butter of banking institutions, into marketable capital market securities. Securitization is the fundamental building block of the shadow banking system.

How the Shadow Banking System Works

In traditional banking, one entity engages in the process of asset transformation—that is, it issues liabilities with one set of desirable characteristics (say, deposits with high liquidity and low risk) to fund the purchase of assets with a different set of characteristics (say, loans with low liquidity and high returns). Securitization, in contrast, is a process of asset transformation that involves a number of different financial institutions working together. These institutions constitute the shadow banking system. In other words, asset transformation accomplished through securitization and the shadow banking system is not done "under one roof," as is traditional banking.

For example, a mortgage broker (more generally referred to as a *loan originator*) will arrange for a residential mortgage loan to be made by a financial institution that will then service the loan (that is, collect the interest and principal payments). This *servicer* then sells the mortgage to another financial institution, which bundles the mortgage with a large number of other residential mortgages. The *bundler* takes the interest and principal payments on the portfolio of mortgages and "passes them through" (pays them out) to third parties. The bundler goes to a *distributor* (typically an investment bank), which designs a security that divides the portfolio of loans into standardized amounts. The distributor then sells the claims to these interest and principal payments as securities, mostly to other financial intermediaries that are also part of the shadow banking system—for example, a money market mutual fund or a pension fund. Schematically, the securitization process is described by the following sequence:

$$\text{Loan origination} \longrightarrow \text{servicing} \longrightarrow \text{bundling} \longrightarrow \text{distribution}$$

Because the securitization process involves the origination of loans and then finally the distribution of securities, securitization is also characterized as an **originate-to-distribute** business model.

At every step of the securitization process, the loan originator, servicer, bundler, and distributor earn a fee. Each of these four institutions has specialized in a particular

element of the financial intermediation process. The shadow banking system, which comprises all of the financial institutions involved in the securitization process, can thus be very profitable if transaction costs and the costs of collecting information are low. Advances in information technology have therefore been critical to the growth of securitization and the shadow banking system. Lower costs of acquiring information make it far easier to sell capital market securities, while lower transaction costs make it cheaper for financial institutions to collect interest and principal payments on bundled loans and then pay them out to the securities holders.

Subprime Mortgage Market A particularly important financial innovation that developed in the 2000s as a result of securitization and the shadow banking system was the **subprime mortgage**, a new class of residential mortgages offered to borrowers with less-than-stellar credit records. Before 2000, only the most creditworthy (prime) borrowers could obtain residential mortgages. Advances in computer technology and new statistical techniques, known as data mining, led to enhanced, quantitative evaluation of the credit risk for residential mortgages. Households with credit records could now be assigned a numerical credit score that predicted how likely the household would be to default on its loan payments. Because it now became easier to assess the risk associated with a pool of subprime mortgages, it became possible to bundle them into a mortgage-backed security, providing a new source of funding for these mortgages. The result was an explosion of subprime mortgage lending, which, as we discuss in Chapter 11, was a key factor that led to the global financial crisis of 2007–2009.

Avoidance of Existing Regulations

The process of financial innovation, as we have discussed it so far, is much like innovation in other areas of the economy: It occurs in response to changes in demand and supply conditions. However, because the financial industry is more heavily regulated than other industries, government regulation is a much greater spur to innovation in this industry. Government regulation leads to financial innovation by creating incentives for firms to skirt regulations that restrict their ability to earn profits. Edward Kane, an economist at Boston College, describes this process of avoiding regulations as "loophole mining." The economic analysis of innovation suggests that when the economic environment changes such that regulatory constraints are so burdensome that large profits can be made by avoiding them, loophole mining and innovation are more likely to occur.

Two sets of regulations have seriously restricted the ability of banks (especially in the United States) to make profits: reserve requirements that force banks to keep a certain fraction of their deposits as reserves (vault cash and deposits in the Federal Reserve System) and restrictions on the interest rates that can be paid on deposits. For the following reasons, these regulations have been major forces behind financial innovation.

1. *Reserve requirements.* Although in Canada reserve requirements were phased out in the Bank of Canada Act in 1992, banks in the United States are still subject to reserve requirements. The key to understanding why reserve requirements led to financial innovation is to recognize that they act, in effect, as a tax on deposits. Because up until 2008 the Federal Reserve did not pay interest on reserves, the opportunity cost of holding them was the interest that a bank could otherwise earn by lending the reserves out. For each dollar of deposits, reserve requirements therefore imposed a cost on the bank equal to the interest rate, i, that could be earned if the reserves could be lent out times the fraction of deposits required as reserves, rd.

The cost $i \times rd$ imposed on the bank was just like a tax on bank deposits of $i \times rd$ per dollar of deposits. This "tax" on deposits rises as interest rates rise.

It is a great tradition to avoid taxes if possible, and banks also play this game. Just as taxpayers look for loopholes to lower their tax bills, banks seek to increase their profits by mining loopholes and by producing financial innovations that enable them to escape the tax on deposits imposed by reserve requirements.

2. *Restrictions on interest paid on deposits.* Canadian banks have never been subject to deposit rate ceilings. In the United States, however, until 1980, legislation prohibited banks in most states from paying interest on chequing account deposits, and through Regulation Q, the Fed set maximum limits on the interest rate that could be paid on saving and time deposits. To this day, banks in the United States are not allowed to pay interest on corporate chequing accounts. The desire of banks to avoid these **deposit rate ceilings** also led to financial innovations.

   If market interest rates rose above the maximum rates that banks paid on savings and time deposits under Regulation Q, depositors withdrew funds from banks to put them into higher-yielding securities. This loss of deposits from the banking system restricted the amount of funds that banks could lend (called **disintermediation**) and thus limited bank profits. Banks had an incentive to get around deposit rate ceilings, because by so doing, they could acquire more funds to make loans and earn higher profits.

We can now look at how the desire to avoid restrictions on interest payments and the tax effect of reserve requirements in the United States led to two important financial innovations.

Money Market Mutual Funds Money market mutual funds issue shares that are redeemable at a fixed price (usually $1) by writing cheques. For example, if you buy 5000 shares for $5000, the money market fund uses these funds to invest in short-term money market securities (Treasury bills, certificates of deposit, commercial paper) that provide you with interest payments. Although money market fund shares effectively function as deposits that earn interest, they are not legally deposits and so are not insured by the Canada Deposit Insurance Corporation (CDIC).

In the United States, money market mutual funds also offer chequing privileges, effectively functioning as chequing account deposits that earn interest. For this reason, money market mutual funds in that country have experienced extraordinary growth since 1970, when they first appeared. Currently, their assets are around US$2.7 trillion. As consumers shift to mutual funds, deposit-taking financial institutions (throughout the world) risk losing a low-cost source of funds.

The first money market mutual fund was created by two Wall Street mavericks, Bruce Bent and Henry Brown, in 1970. In a supreme irony, risky investments by a money market mutual fund founded by Bruce Bent almost brought down the money market mutual fund industry during the global financial crisis in 2008. (See the FYI box "Bruce Bent and the Money Market Mutual Fund Panic of 2008.")

Sweep Accounts Another innovation that enables banks in the United States to avoid the "tax" from reserve requirements is the **sweep account**. In this arrangement, any balances above a certain amount in a corporation's chequing account at the end of a business day are "swept out" of the account and invested in overnight securities that pay interest. Because the "swept out" funds are no longer classified as chequable deposits, they are not subject to reserve requirements and thus are not "taxed." They

FYI **Bruce Bent and the Money Market Mutual Fund Panic of 2008**

Bruce Bent, one of the originators of money market mutual funds, almost brought down the industry during the subprime financial crisis in the fall of 2008. Bent told his shareholders in a letter written in July 2008 that the fund was managed on a basis of "unwavering discipline focused on protecting your principal." He also wrote to the U.S. Securities and Exchange Commission in September 2007: "When I first created the money market fund back in 1970, it was designed with the tenets of safety and liquidity." He added that these principles had "fallen by the wayside as portfolio managers chased the highest yield and compromised the integrity of the money fund." Alas, Bent did not follow his own advice, and his fund, the Reserve Primary Fund, bought risky assets so that its yield was higher than the industry average.

When Lehman Brothers went into bankruptcy on September 15, 2008, the Reserve Primary Fund, with assets of over $60 billion, was caught holding the bag on $785 million of Lehman's debt, which then had to be marked down to zero. The resulting losses meant that on September 16, Bent's fund could no longer afford to redeem its shares at the par value of $1, a situation known as "breaking the buck." Bent's shareholders began to pull their money out of the fund, causing it to lose 90% of its assets.

The fear that this could happen to other money market mutual funds led to a classic panic in which shareholders began to withdraw their funds at an alarming rate. The whole money market mutual fund industry looked like it could come crashing down. To prevent this, the Federal Reserve and the U.S. Treasury rode to the rescue on September 19, 2008. The Federal Reserve set up a facility to make loans to purchase commercial paper from money market mutual funds so they could meet the demands for redemptions from their investors. The U.S. Treasury then put in a temporary guarantee for all money market mutual fund redemptions, and the panic subsided.

Not surprisingly, given the extension of a government safety net to the money market mutual fund industry, there are calls to regulate this industry more heavily. The money market mutual fund industry will never be the same.

also have the advantage that they allow banks in effect to pay interest on these chequing accounts, which otherwise is not allowed under existing regulations. Because sweep accounts have become so popular, they have lowered the amount of required reserves to the degree that most banking institutions do not find reserve requirements binding; in other words, they voluntarily hold more reserves than they are required to.

The financial innovations of sweep accounts and money market mutual funds are particularly interesting because they were stimulated not only by the desire to avoid a costly regulation but also by a change in supply conditions—in this case, information technology. Without low-cost computers to process the additional transactions required by these innovations inexpensively, they would not have been profitable and therefore would not have been developed. Technological factors often combine with other incentives, such as the desire to get around a regulation, to produce innovation.

Financial Innovation and the Decline of Traditional Banking

The traditional financial intermediation role of banking has been to make long-term loans and fund them by issuing short-term deposits, a process of asset transformation commonly referred to as "borrowing short and lending long." Here we examine how financial innovations have created a more competitive environment for the banking

industry, causing the industry to change dramatically, with its traditional banking business going into decline.

To understand why the traditional banking business has declined in size, we need to look at how the financial innovations described earlier have caused banks to suffer declines in their cost advantages in acquiring funds—that is, on the liabilities side of their balance sheet—while at the same time losing income advantages on the assets side of their balance sheet. The simultaneous decline of cost and income advantages has resulted in reduced profitability of traditional banking and an effort by banks to leave this business and engage in new and more profitable activities.

Decline in Cost Advantages in Acquiring Funds (Liabilities) Until 1980, banks were paying low interest rates on chequable deposits. This worked to the banks' advantage because their major source of funds was chequable deposits. Unfortunately, this cost advantage for banks did not last.

The rise in inflation that began in the late 1960s led to higher interest rates, which made investors more sensitive to yield differentials on different assets. The result was the *disintermediation* process, in which people began to take their money out of banks, with their low interest rates on both chequable and savings and time deposits, and began to seek out higher-yielding investments. Also, at the same time, as we have seen, financial innovation led to money market mutual funds, which put the banks at an even further disadvantage because depositors could now obtain chequing-account-like services while earning high interest on their money market mutual fund accounts. One manifestation of these changes in the financial system was that the low-cost source of funds, chequable deposits, declined dramatically in importance for banks.

Decline in Income Advantages on Uses of Funds (Assets) The loss of cost advantages on the liabilities side of the balance sheet for banks is one reason they have become less competitive, but they have also been hit by a decline in income advantages on the assets side from the financial innovations we discussed earlier—junk bonds, securitization, and the rise of the commercial paper market. The resulting loss of income advantages for banks relative to these innovations has resulted in a loss of market share and has led to the growth of the shadow banking system, which has made use of these innovations to enable borrowers to bypass the traditional banking system.

We have seen that improvements in information technology have made it easier for firms to issue securities directly to the public. This change means that instead of going to banks to finance short-term credit needs, many of the banks' best business customers now find it cheaper to go instead to the commercial paper market for funds. The emergence of the junk bond market has also eaten into banks' loan business. Improvements in information technology have made it easier for corporations to sell their bonds to the public directly, thereby bypassing banks. Although well-established companies started taking this route in the 1970s, now lower-quality and less well-known corporate borrowers are using banks less often because they have access to the junk bond market.

We have also seen that improvements in computer technology have led to the growth of the shadow banking system and securitization, whereby illiquid financial assets such as bank loans and mortgages are transformed into marketable securities. Computers enable other financial institutions to originate loans because these institutions can now accurately evaluate credit risk with statistical methods. Computers

also have lowered transaction costs, making it possible to bundle these loans and sell them as securities. When default risk can be easily evaluated with computers, banks no longer have an advantage in making loans. Without their former advantages, banks have lost loan business to other financial institutions even though the banks themselves are involved in the process of securitization. Securitization has been a particular problem for mortgage-issuing institutions, because most residential mortgages are now securitized.

Banks' Responses Canadian banks have sought to maintain adequate profit levels by pursuing new off-balance-sheet activities that are more profitable and in effect embrace the shadow banking system. This strategy, however, has generated concerns about what are proper activities for banks. Nontraditional bank activities can be riskier and, therefore, result in excessive risk taking by banks. Indeed, they led to a substantial weakening of bank balance sheets during the global financial crisis.

The decline of banks' traditional business has driven the banking industry to seek out new lines of business. This search for business opportunities may be beneficial because it may ultimately keep banks vibrant and healthy. Indeed, bank profitability was high up until 2007, and nontraditional, off-balance-sheet activities played an important role in the high bank profits. However, the new directions in banking led to increased risk taking, and thus the decline in traditional banking has required regulators to be more vigilant. Nontraditional bank activities pose new challenges for bank regulators, who, as we saw in Chapter 9, must now be far more concerned about banks' off-balance-sheet activities.

Decline of Traditional Banking in Other Industrialized Countries Forces similar to those in Canada and the United States have led to the decline of traditional banking in other industrialized countries. The loss of banks' monopoly power over depositors has occurred outside Canada and the United States as well. Financial innovation and deregulation are occurring worldwide and have created attractive alternatives for both depositors and borrowers. In Japan, for example, deregulation opened a wide array of new financial instruments to the public, causing a disintermediation process similar to that in Canada and the United States. In European countries, innovations have steadily eroded the barriers that have traditionally protected banks from competition.

Banks in other countries also faced increased competition from the expansion of securities markets and the growth of the shadow banking system. Both financial deregulation and advances in information technology in other countries have improved the availability of information in securities markets, making it easier and less costly for firms to finance their activities by issuing securities rather than going to banks. Further, even in countries where securities markets have not grown, banks have still lost loan business because their best corporate customers have had increasing access to foreign and offshore capital markets, such as the Eurobond market. In smaller economies, like that of Australia, which still do not have well-developed corporate bond or commercial paper markets, banks have lost loan business to international securities markets. In addition, the same forces that drove the securitization process in Canada and the United States have been at work in other countries and have undercut the profitability of traditional banking in these countries as well. Canada and the United States have not been unique in seeing their banks face a more difficult competitive environment. Thus, although the decline of traditional banking occurred earlier in North America than in other countries, the same forces have caused a decline in traditional banking globally.

10.3 STRUCTURE OF THE CANADIAN CHARTERED BANKING INDUSTRY

LO 10.3 Identify the key structural changes in the Canadian commercial banking industry.

As of October 2018, there were 88 banks in Canada, with close to 6000 branches and over 280 000 employees. As Table 10-1 indicates, however, the six largest chartered banks, the Royal Bank of Canada, Canadian Imperial Bank of Commerce (CIBC), Bank of Montreal, the Bank of Nova Scotia (Scotiabank), the Toronto-Dominion Bank, and the National Bank of Canada, together hold about 97% of the assets in the industry.

Schedule I, Schedule II, and Schedule III Banks

The Big Six, together with the Laurentian Bank of Canada, the Canadian Western Bank, and another 25 domestic banks, are Canada's **Schedule I banks**. The remaining 55 banks are **Schedule II banks** (foreign bank subsidiaries, controlled by eligible foreign institutions) and **Schedule III banks** (bank branches of foreign institutions).

Until 1981, foreign banks were not allowed to operate in Canada, and there was no distinction between Schedule I and Schedule II banks. The 1981 revisions to the Bank Act focused on introducing more competition into the Canadian financial services industry. Domestic Canadian banks became Schedule I banks, and subsidiaries of foreign banks became Schedule II banks.

TABLE 10-1 Canadian Banks (as of December 31, 2020)

Bank	Date of Establishment	Assets (in millions) of Bank Assets	Percent (%)
The Big Six			
Royal Bank of Canada	1869	1 746 271	24.86
Canadian Imperial Bank of Commerce	1961	819 159	11.66
Bank of Montreal	1822	992 946	14.14
The Bank of Nova Scotia	1832	1 181 034	16.81
The Toronto-Dominion Bank	1955	1 728 961	24.61
National Bank of Canada	1980	358 393	5.10
Big Six Subtotal		6 826 764	97.19
Laurentian Bank of Canada	1987	44 939	0.64
Canadian Western Bank	1988	34 945	0.50
Eight Domestic Banks Subtotal		6 906 648	98.32
Other Banks			
Domestic Banks		40 258	0.57
Foreign Banks		77 491	1.10
Total All Banks		7 024 397	100.00

Source: OSFI website at www.osfi-bsif.gc.ca. Reprinted with permission.

Schedule I and Schedule II banks have identical powers; the only difference between them is the ownership structure permitted. In particular, according to current ownership policy, all Schedule I banks must be widely held: No individual can own more than 10% of any class of shares. Schedule II banks, however, are exceptions to this rule, if small. In fact, there are three categories of exception. The first exception is that widely held foreign banks can own 100% of a Canadian bank subsidiary. The second exception is that a Schedule II bank may have a significant shareholder (more than 10%) for up to 10 years after chartering, as a transition measure to becoming a Schedule I bank. The third exception, introduced in the 1992 revision of the Bank Act, is that any widely held and regulated Canadian financial institution, other than a bank, may own 100% of a bank. In the case of large Schedule II banks (those with over $5 billion in equity capital), the same widely held ownership rule that applies to Schedule I banks applies.

With the 2001 Bank Act reform (to be discussed in detail later in this chapter), a foreign bank may enter the Canadian banking industry as either a Schedule II or a Schedule III bank. The difference between Schedule II and Schedule III banks is that a Schedule II bank is a Canadian subsidiary of a foreign bank, whereas a Schedule III bank is a foreign bank allowed to branch directly into Canada, under certain restrictions (also to be discussed later in the chapter).

Competition and Technology

Although Canada has a small number of banks relative to other countries, Canadians enjoy one of the most dynamic and competitive financial services industries. Besides chartered banks, there are over 4000 financial institutions providing financial services. These include trust and loan companies, credit unions and *caisses populaires*, government savings institutions, insurance companies, pension funds, mutual funds, and investment dealers.

New technology and the Internet have also helped in the development of a more competitive and innovative banking system in Canada. They have enabled new entrants to come into the Canadian financial services market and provide innovative services that meet the needs of modern consumers. For example, in November 2012, Scotiabank completed its acquisition of ING Bank of Canada, a former Canadian banking subsidiary of a major Netherlands banking and insurance conglomerate, and changed its name to Tangerine. Tangerine, like Vancity Community Investment Bank (formerly known as Citizens Bank of Canada), are mostly virtual banks (without bricks-and-mortar branches) offering an array of banking services on the Internet.

10.4 COMPARISON WITH THE UNITED STATES

LO 10.4 Identify the key differences between the Canadian and U.S. banking industries.

The structure of the commercial banking industry in Canada, although similar to that in many other industrialized countries, is radically different from that in the United States. There are approximately 6000 commercial banks in the United States, far more than in any other country in the world. Japan, for example, has fewer than 100 commercial banks—a mere fraction of the number of banks in the United States, even though its economy and population are half the size of those of the United States—and, as already noted, Canada has 88 banks.

As Table 10-2 indicates, the United States has an extraordinary number of small banks; over 22% of the banks have less than US$100 million in assets, and the 10 largest commercial bank in the United States (listed in Table 10-3) together hold under 70% of the assets in their industry.

Most industries in the United States have far fewer firms than the commercial banking industry; typically, large firms tend to dominate these industries to a greater extent than in the commercial banking industry. (Consider the computer software industry, which is dominated by Microsoft, or the automobile industry, which is dominated by General Motors, Ford, Daimler-Chrysler, Toyota, and Honda.) Does the large number

TABLE 10-2 Size Distribution of FDIC-Insured Banks, March 31, 2020

Assets	Number of Banks	Share of Banks (%)	Share of Assets Held (%)
Less than $100 million	1124	22.0	0.5
$100 million–$1 billion	3168	61.9	6.7
$1 billion–$10 billion	680	13.3	10.0
$10 billion–$250 billion	131	2.6	32.3
More than $250 billion	13	0.3	50.4
Total	5116	100.00	100.00

Source: FDIC Quarterly Banking Profile, https://www.fdic.gov/bank/analytical/qbp/index.html.

TABLE 10-3 Ten Largest U.S. Banks, 2020

Bank	Assets ($ trillions)	Share of All Commercial Bank Assets (%)
1. J.P. Morgan Chase & Co.	2.74	15.4
2. Bank of America Corp.	2.38	13.4
3. Citigroup Inc.	1.96	11.0
4. Wells Fargo & Co.	1.89	10.6
5. Goldman Sachs Group	0.93	5.2
6. Morgan Stanley	0.88	4.9
7. U.S. Bankcorp	0.48	2.7
8. PNC Financial Services Group	0.39	2.2
9. TD Group US	0.38	2.1
10. Capital One Financial Corp.	0.37	2.1
Total	12.40	69.6

Source: From Bankrate.com—Compare mortgage, refinance, insurance, CD rates: http://www.bankrate.com/banking/americas-top-10-biggest-banks/#slide=1.

of banks in the commercial banking industry and the absence of a few dominant firms suggest that commercial banking is more competitive than other industries?

Restrictions on Branching

The presence of so many commercial banks in the United States actually reflects past regulations that restricted the ability of these financial institutions to open branches. Each state had its own regulations on the type and number of branches that a bank could open. Regulations on both coasts, for example, tended to allow banks to open branches throughout a state; in the middle part of the country, regulations on branching were more restrictive, in some cases allowing no branches at all. The McFadden Act of 1927, which was designed to put national banks and state banks on equal footing, effectively prohibited banks from branching across state lines and forced all national banks to conform to the branching regulations of the state in which their headquarters were located.

The McFadden Act and state branching regulations constituted strong anticompetitive forces in the U.S. commercial banking industry, allowing many small banks to stay in existence because larger banks were prevented from opening branches nearby. If competition is beneficial to society, why did regulations restricting branching arise in the United States? The simplest explanation is that the American public has historically been hostile to large banks. States with the most restrictive branching regulations were typically ones in which populist antibank sentiment was strongest in the nineteenth century. (These states usually had large farming populations whose relations with banks periodically became tempestuous when banks would foreclose on farmers who couldn't pay their debts.) The legacy of nineteenth-century politics was a banking system with restrictive branching regulations and hence an inordinate number of small banks. However, as we will see later in this chapter, branching restrictions have been eliminated, and the United States has moved toward nationwide banking.

Response to Branching Restrictions

An important feature of the U.S. banking industry is that competition can be repressed by regulation but not completely quashed. As we saw earlier in the chapter, the existence of restrictive regulation stimulates financial innovations that get around these regulations in the banks' search for profits. Regulations restricting branching have stimulated similar economic forces and have promoted the development of two financial innovations: bank holding companies and automated teller machines.

Bank Holding Companies A holding company is a corporation that owns several different companies. This form of corporate ownership has important advantages for banks. It has allowed them to circumvent restrictive branching regulations, because the holding company can own a controlling interest in several banks even if branching is not permitted. Furthermore, a bank holding company can engage in other activities related to banking, such as the provision of investment advice, data processing and transmission services, leasing, credit card services, and servicing of loans in other states.

The growth of bank holding companies has been dramatic over the past three decades. Today bank holding companies own almost all large banks, and more than 90% of all commercial bank deposits are held in banks owned by holding companies.

Automated Teller Machines Another financial innovation that resulted from the desire to avoid restrictions on branching is the ATM. Banks quickly realized that if they did not own or rent the ATM, but instead let it be owned by someone else and paid a fee for each transaction, the ATM would probably not be considered a branch of the bank and thus would not be subject to branching regulations. This is exactly what the regulatory agencies and courts in most states in the United States concluded. Because they enable banks to widen their markets, a number of these shared facilities (such as Cirrus and NYCE) have been established nationwide. Furthermore, even when an ATM is owned by a bank, states typically have special provisions that allow wider establishment of ATMs than is permissible for traditional "brick-and-mortar" branches.

As we saw earlier in this chapter, avoiding regulation was not the only reason for the development of the ATM. The advent of cheaper computer and telecommunications technology enabled banks to provide ATMs at low cost, making them a profitable innovation. This example further illustrates that technological factors often combine with incentives, such as the desire to avoid restrictive regulations, to produce financial innovation.

10.5 COMPETITION ACROSS ALL FOUR PILLARS

LO 10.5 Summarize the factors that led to consolidation in the commercial banking industry.

Another important feature of the structure of the banking industry in Canada until recently was the separation of the banking and other financial services industries—such as securities, insurance, and real estate. Regulations enforced the separation of institutions according to their core financial service, and only four distinct types of financial services were identified: banking, brokerage, trusts, and insurance. This approach to regulation by institution (versus regulation by function) has been known as the **four-pillar approach**. The separation of the four pillars prohibited chartered banks from engaging in insurance and real estate activities. In turn, it prevented investment banks and insurance companies from engaging in commercial banking activities and thus protected banks from competition.

Convergence

In recent years, however, financial markets have opened up, and Canada's traditional four-pillar system has changed. Despite the prohibitions in the legislation, the pursuit of profits and financial innovation stimulated both banks and other financial institutions to bypass the intent of the legislation and encroach on each other's traditional territory. For example, credit unions have long offered insurance to their members, and brokerage firms have engaged in the traditional banking business of issuing deposit instruments with the development of money market mutual funds and cash management accounts.

Not surprisingly, the regulatory barriers between banking and other financial services markets have been coming down in response to these forces. Before the 1950s, for example, legislation allowed chartered banks to make loans for commercial purposes only and prohibited them from making residential mortgage loans. It was only after the 1967 revision of the Bank Act that banks were allowed to make conventional

residential mortgage loans, thereby directly competing with trust and loan companies and credit unions and *caisses populaires*.

In the 1980s, the Bank Act was amended to allow Canadian and foreign financial institutions to own up to 100% of securities firms. Moreover, the 1990s revisions to the Bank Act allowed cross-ownership via subsidiaries between financial institutions. Chartered banks, for example, can either buy independent investment dealers or expand on their own into capital raising, brokerage, and other securities activities. As a result, the Big Six now dominate the investment-banking industry through their investment brokerage subsidiaries. As a result of these legislative changes, Canada's traditional four pillars of financial services—banking, brokerage, trusts, and insurance—have converged into a single financial services marketplace.

Not surprisingly, the advent of the Web and improved computer technology is another factor driving bank consolidation. Economies of scale have increased because large, upfront investments are required to set up information technology platforms for financial institutions. To take advantage of these economies of scale, banks have needed to get bigger, and this development has led to additional consolidation. Information technology has also been increasing **economies of scope**, the ability to use one resource to provide many different products and services. For example, details about the quality and creditworthiness of firms not only inform decisions about whether to make loans to them but also can be useful in determining at what price their shares should trade. Similarly, once you have marketed one financial product to an investor, you probably know how to market another. Businesspeople describe economies of scope by saying that there are "synergies" between different lines of business, and information technology is making these synergies more likely. The result is that consolidation is taking place not only to make financial institutions bigger but also to increase the combination of products and services they can provide.

This consolidation has had two consequences. First, different types of financial intermediaries are encroaching on each other's territory, making them more alike. Second, consolidation has led to the development of large, complex banking organizations.

Implications for Financial Consolidation

As we have seen, recent legislation has stimulated consolidation of the banking industry. The financial consolidation process will speed up even faster in the future, because the way is now open to consolidation in terms not only of the number of banking institutions but also across financial service activities. Given that information technology is increasing economies of scope, mergers of banks with other financial service firms like that of Citicorp and Travelers in the United States should become increasingly common, and more mega-mergers are likely to be on the way. Banking institutions are becoming not only larger but also increasingly complex organizations, engaging in the full gamut of financial service activities. The trend toward larger and more complex banking organizations has been accelerated by the global financial crisis of 2007–2009. (See the FYI box "The Global Financial Crisis and the Demise of Large, Free-Standing Investment Banks.")

Separation of Banking and Other Financial Services Industries Throughout the World

In the aftermath of the Great Depression, not many other countries followed the lead of Canada and the United States in separating the banking from other financial services industries. In fact, in the past this separation was the most prominent difference between

FYI The Global Financial Crisis and the Demise of Large, Free-Standing Investment Banks

Although the move toward bringing financial service activities into larger, more complex banking organizations was inevitable after the demise of **Glass-Steagall** (an Act passed in 1933 by the U.S. Congress to legally separate investment banking from commercial banking), no one expected it to occur as rapidly as it did in 2008. Over a six-month period from March to September 2008, all five of the largest free-standing investment banks in the United States ceased to exist in their old form. When Bear Stearns the fifth-largest investment bank, revealed its large losses from investments in subprime mortgage securities, it had to be bailed out by the U.S. Federal Reserve in March 2008; the price it paid was a forced sale to J.P. Morgan for less than one-tenth of what it had been worth only a year or so before. The Bear Stearns bailout made it clear that the government safety net had been extended to investment banks. The trade-off is that investment banks will be subject to more regulation, along the lines of commercial banks, in the future.

Next to go was Lehman Brothers, the fourth-largest investment bank in the United States, which declared bankruptcy on September 15, 2008. Only one day before, Merrill Lynch, the third-largest investment bank, which also suffered large losses on its holdings of subprime securities, announced its sale to Bank of America for less than half of its year-earlier price. Within a week, Goldman Sachs and Morgan Stanley, the largest and second-largest investment banks, both of which had less exposure to subprime securities, nevertheless saw the writing on the wall. They realized that they would soon become regulated on a similar basis and decided to become bank holding companies so that they could access insured deposits, a more stable funding base.

It was the end of an era. Large, free-standing investment-banking firms in the United States are now a thing of the past.

banking regulation in Canada and the United States and that in other countries. Around the world, there are three basic frameworks for the banking and securities industries.

The first framework is *universal banking*, which exists in Germany, the Netherlands, and Switzerland. It provides no separation at all between the banking and securities industries. In a universal banking system, commercial banks provide a full range of banking, securities, real estate, and insurance services, all within a single legal entity. Banks are allowed to own sizable equity shares in commercial firms, and often they do.

The *British-style universal banking system*, the second framework, is found in the United Kingdom and countries with close ties to it, such as Canada and Australia, and now the United States and Japan. The British-style universal bank engages in securities underwriting, but it differs from the German-style universal bank in three ways: Separate legal subsidiaries are more common, bank equity holdings of commercial firms are less common, and combinations of banking and insurance firms are less common.

The third framework, which existed in Japan, features some legal separation of the banking and other financial services industries. A major difference between the British-style and Japanese banking systems is that Japanese banks are allowed to hold substantial equity stakes in commercial firms, whereas British-style universal banks are not. In addition, most American banks use a bank-holding-company structure, but bank holding companies are illegal in Japan. Although the banking and securities industries are legally separated in Japan under Section 65 of the Japanese Securities Act, commercial banks are increasingly being allowed to engage in securities activities and are thus becoming more like British-style universal banks.

10.6 THE NEAR BANKS: REGULATION AND STRUCTURE

LO 10.6 Summarize the distinctions between chartered banks and trust and loan companies and credit unions and *caisses populaires*.

Not surprisingly, the regulation and structure of the near banks (trust and loan companies and credit unions and *caisses populaires*) closely parallel the regulation and structure of the chartered banking industry.

Trust Companies

Over the years, the Bank Acts have denied chartered banks the power to function as corporate **trustees** (or fiduciaries). Unlike the situation in the United States, legislators in Canada reasoned that deposit-taking financial institutions might face a conflict of interest if they were to act as both financial fiduciaries and banks. So trust companies were established, under a variety of provincial and federal laws, and specialized in the provision of fiduciary services. As financial fiduciaries, trust companies administer estates, trusts, and agencies (i.e., assets that belong to someone else) for a fee and under conditions prescribed in a contract.

Over the years, the structure of the trust industry has changed significantly, and today trusts offer similar services as those offered by banks (i.e., deposit taking, loans) as well as fiduciary services. Moreover, over the years the Bank Acts have allowed regulated federal financial institutions (domestic chartered banks and life insurance companies) to own trust companies, thereby providing fiduciary services indirectly through their subsidiaries. As a result, and with the acquisition of Canada Trust (Canada's largest trust company) by the Toronto-Dominion Bank in early 2000, trust companies now constitute a relatively small market segment.

Loan Companies

The development of trust companies was paralleled by the growth of loan companies. The concept of loan companies came from the building societies in the United Kingdom (during the early part of the nineteenth century), whose purpose was to enable members to acquire land, build homes, or develop farms. Today's loan companies take deposits and primarily make residential mortgage loans. They do not act as trustees unless they are licensed specifically for that purpose. Over the years, the loan companies together with the trust companies formed the second pillar of the traditional financial services industry in Canada. However, as you saw in this chapter, financial innovation, competition, and regulatory evolution have significantly changed the competitive position of financial institutions in recent years.

Trust and loan companies operate under a charter issued by either the federal government or one of the provincial governments. Federally incorporated trust and loan companies come under the federal Trust and Loan Companies Act and are regulated and supervised by the Office of the Superintendent of Financial Institutions (OSFI). They must also register in all the provinces in which they operate and must conform to the regulations of those provinces. In the case of trust companies, the fiduciary component of their business is subject only to provincial legislation, even if the company is federally incorporated. Deposit insurance for trust and loan companies outside Québec is provided by the CDIC (up to $100 000 per account). The Québec Deposit Insurance Board insures deposits for Québec trust and loan companies on terms similar to the CDIC's.

Trust and loan companies are funded almost entirely by chequable and non-chequable savings deposits, term deposits, guaranteed investment certificates, and debentures; together, they account for about 85% of the balance sheet. Their risk asset portfolio is made up mostly of residential mortgages and loans; together they account for about 30% of assets. The low-risk assets are largely deposits with regulated financial institutions.

Cooperatives: Credit Unions and *Caisses Populaires*

Cooperatives are small lending institutions organized around a particular group of individuals with a common bond (union members or employees of a particular firm). Alphonse Desjardins formed Canada's first cooperative bank in 1900 in Québec, and it was based on the cooperative movements in Europe, which, among other things, stressed the provision of credit to the "little man." Today, there are two cooperative financial systems in Canada: the *caisses populaires* system in Québec and the credit union system in other parts of the country.

Credit unions and *caisses populaires* offer products and services ranging from mortgages and car loans to credit lines and automated teller machines. Because their members share a common goal, credit unions and *caisses populaires* are typically quite small; most are about the size of a single bank branch and hold less than $10 million of assets, with the largest being Vancity, with assets around $23 billion and over half a million members. The credit unions and *caisses populaires* are established under provincial legislation and are non-profit-seeking financial institutions. Unlike the shareholder-owned chartered banks that accept deposits from and lend funds to the general public, the cooperative banks accept deposits (shares) from and lend funds only to their members. Members have voting rights and elect a board of directors, which determines the union's lending and investment policies.

As member-owned, independent financial firms, credit unions and *caisses populaires* constitute an alternative financial system, different from the profit-seeking banking system. They have also developed their own set of institutions, including central banking and deposit insurance arrangements. In particular, each province has a central credit union, owned by the member credit unions, providing financial services to individual credit unions. All central credit unions outside Québec are members of the Canadian Credit Union Association, which serves as the third tier for the credit union movement; it coordinates various functions and provides cheque-clearing services for all provincial central credit unions.

In Québec, *caisses populaires* are assembled under a single federation, the Fédération des caisses Desjardins du Québec, which has a similar structure to the provincial central credit unions in the rest of Canada but with considerably broader regulatory responsibilities. The federation dominates retail and corporate financial intermediation in Québec. It owns a major P&C (property and casualty) insurance group, Desjardins General Insurance; a life and health insurance company, Desjardins Financial Security; a venture capital firm, Desjardins Venture Capital; and an investment dealer and discount brokerage, Desjardins Securities.

Credit unions and *caisses populaires* are not directly covered by the CDIC. However, each provincial government has an agency, commonly called a stabilization fund, that has a line of credit with the provincial Treasury and provides deposit guarantees for credit unions. Deposits in New Brunswick and Prince Edward Island are insured up to $60 000 per account; in Ontario and British Columbia up to $100 000; in Nova Scotia and Newfoundland and Labrador up to $250 000; and in Alberta, Saskatchewan,

and Manitoba there are no limits on the amount of coverage provided. In Québec, the Québec Deposit Insurance Board, the same provincial government agency that insures deposits in other deposit-taking financial institutions in Québec, also provides deposit guarantees for *caisses populaires* on terms similar to the CDIC's.

Government Savings Institutions

In addition to the near banks (trust and loan companies and credit unions and *caisses populaires*), there are some government-operated deposit-taking institutions, such as Alberta Treasury Branches (ATB) Financial.

The province of Alberta established Treasury branches back in 1938 in response to Albertans' needs in remote areas. Today, there are over 170 branches in over 240 communities across the province operating in three target markets: individual financial services, agricultural operations, and independent business. ATB Financial (with assets in excess of $55 billion) is funded almost entirely by demand, notice, and fixed-term deposits, and its risk asset portfolio is made up largely of residential mortgages and personal, commercial, and agricultural loans.

10.7 INTERNATIONAL BANKING

LO 10.7 Identify the reasons for Canadian banks to operate in foreign countries and for foreign banks to operate in Canada.

Canadian banks have a well-developed presence in the global financial services marketplace, which varies among the individual institutions. From its inception, for example, the Bank of Montreal found some of its best opportunities in international operations and was soon joined by the Canadian Imperial Bank of Commerce and the Bank of Nova Scotia. The spectacular growth in international banking can be explained by three factors.

First is the rapid growth in international trade and multinational (worldwide) corporations that has occurred in recent years. When Canadian firms operate abroad, they need banking services in foreign countries to help finance international trade. For example, they might need a loan in a foreign currency to operate a factory abroad. And when they sell goods abroad, they need to have a bank exchange the foreign currency they have received for their goods into Canadian dollars. Although these firms could use foreign banks to provide them with these international banking services, many of them prefer to do business with the Canadian banks with which they have established long-term relationships and that understand Canadian business customs and practices. As international trade has grown, international banking has grown with it.

Second, Canadian banks have been able to earn substantial profits by being very active in global investment banking, in which they underwrite foreign securities. They also sell insurance abroad, and they derive substantial profits from these investment-banking and insurance activities.

Third, Canadian banks have wanted to tap into the large pool of Eurocurrencies—currencies deposited in banks outside the home country. To understand the structure of Canadian banking overseas, let us first look at the Eurocurrencies market, an important source for international banking.

Global **Ironic Birth of the Eurodollar Market**

One of capitalism's great ironies is that the Eurodollar market, one of the most important financial markets used by capitalists, was created by the Soviet Union. In the early 1950s, during the height of the Cold War, the Soviets had accumulated a substantial amount of dollar balances held by banks in the United States. Because the Russians feared that the U.S. government might freeze these assets in the United States, they wanted to move the deposits to Europe, where they would be safe from expropriation. (This fear was not unjustified—consider the U.S. freeze on Iranian assets in 1979 and on Iraqi assets in 1990.) However, they also wanted to keep the deposits in dollars so that they could be used in their international transactions. The solution to the problem was to transfer the deposits to European banks but to keep the deposits denominated in dollars. When the Soviets did this, the Eurodollar was born.

Eurocurrencies Market

The most important of the Eurocurrencies are Eurodollars, which originated after World War II with U.S.-dollar deposits in European banks. They were created when deposits in accounts in the United States were transferred to a bank outside the United States and were kept in the form of U.S. dollars. (For a discussion of the birth of the Eurodollar, see the Global box "Ironic Birth of the Eurodollar Market.") For example, if Rolls-Royce PLC deposits a US$1 million cheque, written on an account at an American bank, in its bank in London—specifying that the deposit be payable in U.S. dollars—$1 million in Eurodollars is created.[3] More than 90% of Eurodollar deposits are time deposits, more than half of which are certificates of deposit with maturities of 30 days or more. The Eurodollar market is massive, over US$10 trillion, making it one of the most important financial markets in the world economy.

Although most offshore deposits are denominated in U.S. dollars, some are also denominated in other currencies. Collectively, these offshore deposits are referred to as Eurocurrencies. A Canadian dollar–denominated deposit held in London, for example, is called a Euro Canadian dollar, and a Japanese yen–denominated deposit held in London is called a Euroyen. Why would companies such as Rolls-Royce want to hold Eurocurrencies? First, some currencies are widely used in international trade, so Rolls-Royce might want to hold deposits in these currencies to conduct its international transactions. Second, Eurocurrencies are "offshore" deposits—they are held in countries that will not subject them to regulations such as reserve requirements or restrictions (called *capital controls*) on taking the deposits outside the country.

The main centre of the Eurocurrencies market is London, a major international financial centre for hundreds of years. Eurocurrencies are also held outside Europe in locations that provide offshore status to these deposits—for example, Singapore, the Bahamas, and the Cayman Islands.

The minimum transaction in the Eurocurrencies market is typically $1 million, and approximately 75% of Eurocurrency deposits are held by banks. Plainly, you and I

[3]Note that the bank in London keeps the $1 million on deposit at the American bank, so the creation of Eurodollars has not caused a reduction in the amount of bank deposits in the United States.

are unlikely to come into direct contact with Eurocurrencies. The Eurocurrencies market is, however, an important source of funds to Canadian banks. Rather than using an intermediary and borrowing all the deposits from foreign banks, Canadian banks decided that they could earn higher profits by opening their own branches abroad to attract these deposits. Consequently, the Eurocurrencies market has been an important stimulus to Canadian banking overseas.

Canadian Banking Overseas

Canadian banks have been present in international financial markets for over 100 years, providing services to Canadians and multinational businesses. The international presence of the Big Six varies among the individual institutions. In particular, the Bank of Montreal, the Canadian Imperial Bank of Commerce, and the Toronto-Dominion Bank have a significant presence in the United States, whereas Scotiabank has established a presence in South America and the Royal Bank in Europe and Asia.

During the 1970s and early 1980s, a large proportion of the banks' foreign lending was in **sovereign loans**—loans to foreign governments and their agencies in the less developed countries, particularly Mexico, Brazil, Venezuela, Argentina, and Chile. Most of this activity in international lending was unregulated, with near-disastrous consequences. One example is the international debt crisis, which had its origin in the oil price shocks of the 1970s. In particular, the 1973–1974 increase in the price of oil was a bonanza for some oil-exporting countries like Mexico but a disaster for oil-importing countries like Brazil, which had to either cut their living standards or borrow massively abroad in order to pay their higher oil bills. At the time, real interest rates were very low (in fact, negative), and the oil importers couldn't resist the temptation to borrow abroad.

At the same time, the oil exporters were depositing huge sums in banks, and as a result the banks were lending not only to the oil-importing countries but also to the oil-exporting countries, because the latter had large oil reserves and seemed like good credit risks. The banks underestimated the **indebtedness** of these countries—the total amount these countries had borrowed from banks—and, as a result, were severely punished in the early 1980s when the recession hit and real interest rates increased significantly. With Argentina, Brazil, Mexico, and Peru threatening to default on their loans, the banks had two choices: to reschedule their loans and make more loans to these countries (to enable them to pay the interest on the debt) or to declare these countries in default and acknowledge large losses on their balance sheets. The banks chose to make more loans, because in many cases the losses would have been large enough to destroy them.

As a result of their lending experience in Latin America, the Big Six have withdrawn from certain countries and focused more of their international activities in the United States. Moreover, the international activities of Canadian banking organizations are now regulated primarily by OSFI.

Foreign Banks in Canada

The growth in international trade has not only encouraged Canadian banks to open offices overseas but has also encouraged foreign banks to establish offices in Canada. Foreign banks have been extremely successful in Canada. Over the past 35 years, since the 1981 revision to the Bank Act, globally prominent foreign banks have set up and expanded banking subsidiaries in Canada. Foreign banks are a highly fragmented group and currently hold about 7% of total Canadian bank assets, with HSBC Bank Canada (a subsidiary of HSBC Holdings PLC) being the seventh-largest bank in Canada, with

offices in every province except Prince Edward Island. It should be noted, however, that these institutions target specific groups, achieving a higher representation within their target groups than their national share would suggest. For example, HSBC, the largest of the Schedule II banks, enjoys a strong presence and success in the Chinese communities of British Columbia and Ontario.

Foreign banks may enter the Canadian financial services industry as either Schedule II or Schedule III banks. As already noted, Schedule II banks don't have to be widely held if small. If, however, their equity capital exceeds $1 billion, then at least 35% of it must be widely held. In the case that their equity capital exceeds $5 billion, then the same widely held ownership rule applies as for Schedule I banks. The major difference between Schedule II and Schedule III banks is that Schedule III banks can branch directly into Canada, following authorization by the minister of finance, whereas Schedule II banks can add branches to their initial branch only with ministerial approval. However, Schedule III banks cannot take retail deposits (i.e., deposits less than $150 000) and, as a result, have the advantage of not being subject to regulations that apply to full-service banks (such as requirements for CDIC insurance). Given that most Schedule II banks do little retail deposit gathering, it is likely that in the future many Schedule II banks will become Schedule III banks.

The internationalization of banking, both by Canadian banks going abroad and by foreign banks entering Canada, has meant that financial markets throughout the world have become more integrated. As a result, there is a growing trend toward international coordination of bank regulation, one example of which is the Basel Accords, discussed in Chapter 9, which were established to standardize minimum bank capital requirements in industrialized countries. Financial market integration has also encouraged bank consolidation abroad, culminating in the creation of the first trillion-dollar bank with the merger of the Industrial Bank of Japan, Dai-Ichi Kangyo Bank, and Fuji Bank, in 2002. Another development has been the importance of foreign banks in international banking as shown in Table 10-4 for 2020. The implications of this financial market integration for the operation of our economy are examined further in Chapter 19 when we discuss the international financial system in more detail.

TABLE 10-4 Ten Largest Banks in the World, 2020

Bank	Assets (US$ trillions)
1. Industrial and Commercial Bank of China, China	4.32
2. China Construction Bank Corp., China	3.82
3. Agricultural Bank of China, China	3.70
4. Bank of China, China	3.39
5. JP Morgan Chase, U.S.	3.14
6. HSBC Holdings PLC, United Kingdom	2.92
7. Mitsubishi UFJ Financial Group, Japan	2.89
8. Bank of America, U.S.	2.62
9. BNP Paribas, France	2.43
10. Credit Agricole Group, France	1.98

Source: FXSSI: https://fxssi.com/top-20-largest-world-banks-in-current-year.

10.8 THE 2001 BANK ACT REFORM

LO 10.8 Assess the reasons for separating banking from other financial services through legislation.

We have seen that fundamental structural changes make possible new financial products and services, increasing the competitive environment in the financial services industry and changing the financial intermediation role of banking at Internet speed. It is against this backdrop of phenomenal change that the federal government introduced legislation reforming the policy framework of the Canadian financial services sector.[4] This legislation took effect in October 2001 and is known as **Bank Act reform**.

In what follows, we provide an overview of the key elements of the new regulatory environment for financial institutions that has the potential to dramatically change the face of competition in Canada's financial services marketplace. As you will see, the new laws establish the regulatory framework to accelerate changes already taking place throughout the Canadian and world economies and introduce new opportunities for strategic alliances and partnerships, with the objective of fostering more competition and providing more innovative products and services to Canadians.

Bank Holding Companies

Before Bank Act reform, the organizational structure of Canada's bank financial groups was based on the "bank-as-parent" model, where all banking functions and all subsidiaries of the bank are subject to the same regulation. Under the new legislation, bank financial groups have the option of organizing themselves under a holding company.[5] A holding company is a corporation that owns several different companies. For example, under a holding company structure, a bank financial group may have a banking subsidiary, an insurance subsidiary, a securities subsidiary, and another subsidiary for its unregulated businesses.

Most developed countries permit holding company structures, and the growth of holding companies has been dramatic over the past three decades. Today, in the United States, for example, holding companies own almost all large banks, and more than 90% of all commercial bank deposits are held in banks owned by holding companies. The holding company form of corporate ownership has important advantages for bank financial groups in that (1) it allows them to engage in other activities related to banking, such as the provision of investment advice, data processing and transmission services, leasing, and credit card services; (2) a holding company structure allows for lighter regulation throughout the bank financial group because certain activities (those not involving retail deposit taking and insurance) can be undertaken by less regulated, non-bank-affiliated companies held by the holding company parent rather than the regulated operating bank; and (3) the holding company model provides bank financial groups increased flexibility to achieve economies of scale and scope through strategic partnerships, alliances, and joint ventures.

[4]Since 1992, the practice of reviewing legislation governing Canada's chartered banks has been extended to reviewing legislation governing all federal financial institutions.

[5]The new legislation introduced a holding company regime for Canada's insurance companies as well. Under the legislation, insurance holding companies are required to have an investment in at least one life insurance company and are regulated under the Insurance Companies Act. Bank holding companies are regulated under the Bank Act and are required to have an investment in at least one bank.

The big advantage of the new regime for establishing holding companies is that financial groups will be able to move parts of their heavily regulated business into less regulated affiliates under a common holding company. The holding company would be a viable option for financial groups if the transition to a holding company would be tax-neutral and without increased costs or regulatory burdens. If, for example, additional taxes or heavier regulation were to arise as a result of a restructuring into a holding company, then the holding company option wouldn't be feasible. For this reason, the new legislation also provides a set of transitional rules to address unintended cost consequences that would be triggered by the transition to a holding company.

Permitted Investments

A second item of the 2001 legislation pertains to the type of investments federally regulated financial institutions are allowed to make. At present, there is a restrictive list of activities beyond banking that banks, for example, can get involved in. The new legislation, however, provides greater flexibility for bank involvement in the information technology area (and, in particular, the Internet and wireless technology). It permits bank financial groups to establish and operate information services entities utilizing recent advances in Internet and wireless banking and voice recognition technologies.

New information technologies are critical to the ability of Canadian banks and insurance companies to provide new financial products and services and adapt to the changing marketplace. Although bank and insurance company involvement in the information technology area is subject to regulation, the new permitted investment regime will enhance the ability of banks and insurance companies to pursue strategic alliances and joint ventures and will further accelerate the technological advances that are already taking place and revolutionize the financial services sector.

Ownership Rules

A third policy measure that can significantly change the face of competition in Canada's financial services marketplace is the new ownership regime that enables investors to take a greater equity interest in widely held bank financial groups by expanding the definition of "widely held." In particular, the 2001 legislation increased the limit (in place since 1967) a single shareholder can own of a widely held financial institution (either a bank holding company or a bank subsidiary under the holding company) from 10% of any class of shares to 20% of voting shares and 30% of nonvoting shares. The new legislation, however, does not permit a single shareholder to own more than 10% of both a bank holding company and a bank subsidiary under the holding company at the same time. Moreover, acquisitions of more than 10% are subject to approval by the minister of finance based on a "fit and proper person" test.

The new legislation also includes a *three-tiered ownership regime*. Small banks (those with equity capital under $1 billion) don't have to be widely held and can be wholly owned (i.e., have one particular investor own 100% of the shares).

Medium-sized banks (and bank holding companies) with shareholders' equity between $1 billion and $5 billion can be closely held provided that there is a 35% public float (that is, they could have a single shareholder own up to 65% of their shares). Large banks (and bank holding companies), those with shareholders' equity in excess of $5 billion, are required to be widely held.

The new ownership regime, together with other provisions in the legislation, such as the lowering of the capital needed to create a bank from $10 million to $5 million

and the allowance of domestic and foreign commercial enterprises (such as department stores and grocery chains) to establish small and medium-sized banks, are fundamentally changing Canada's financial sector.

The CP Act and Access to the Payments and Clearance System

Before the 2001 financial-sector legislation, membership in Payments Canada (then the Canadian Payments Association, a nonprofit organization formed in 1980 by an Act of Parliament to operate Canada's payments systems) was limited to the Bank of Canada and the deposit-taking financial institutions—chartered banks, trust and loan companies, and credit unions and *caisses populaires*. The 2001 legislation introduced some important changes for the Canadian Payments Association and also renamed the Canadian Payments Association Act to Canadian Payments Act (CP Act).

In particular, the CP Act extends eligibility for membership in Payments Canada—and therefore access to Canada's two domestic payments systems, the Large Value Transfer System and the Automated Clearing Settlement System, both to be discussed in detail in Chapter 16—to non-deposit-taking financial institutions, such as life insurance companies, securities dealers, and money market mutual funds. This regulatory change will significantly affect Canada's financial services sector, since it will allow these organizations to provide bank-like services, such as chequing accounts and debit cards, without being banks, thereby directly competing with banks, trust and loan companies, and credit unions and *caisses populaires*.

Expanding access to the payments and clearance system, by allowing non-deposit-taking financial institutions to participate, will further accelerate the process of the blurring of distinction between deposit-taking and non-deposit-taking financial institutions. As already noted, this process started in 1987, when securities dealers were allowed to own banks, and was reinforced by the 1992 federal financial reforms that permitted cross-ownership of financial institutions.

Merger Review Policy

The government has also issued a statement establishing a process for reviewing mergers involving large banks—banks like the Bank of Montreal and CIBC—with shareholder equity in excess of $5 billion. By doing so, the government acknowledges that mergers are a legitimate business option that should be available to Canadian bank financial groups. The bank merger review process, however, unlike those in other countries such as the United States and the United Kingdom, is political, having Parliament directly involved in it.

Under the new merger review policy, the merger partners are required to submit a public interest impact assessment covering various effects of the merger, such as the impact on the structure and competition of the banking industry, branch closures, and job losses. The merger proposal would then be submitted to the House of Commons Standing Committee on Finance and the Standing Senate Committee on Banking, Trade, and Commerce for consideration and public hearings. Each of these committees would report to the minister of finance, who would make these reports public, together with a report from the Competition Bureau on the competitive aspects of the proposed merger and a report from OSFI on the prudential aspects of the proposed merger. The minister of finance would decide whether the proposed merger would be allowed to proceed.

The government has also indicated that it will not allow mergers between large banks and large demutualized life insurance companies such as Manulife and Sun Life. We would note, however, that in other countries, such as Australia, Germany,

the Netherlands, Switzerland, the United Kingdom, and the United States, mergers of banks, insurance companies, and other financial services providers are not prohibited.

The National Financial Services OmbudService

Another significant achievement is the creation of a National Financial Services OmbudService (NFSO), which began operations on July 1, 2002. This service has been created, with the support of the federal and provincial governments, by the banking sector (through the Canadian Bankers Association), the insurance sector (through the Insurance Bureau of Canada and the Canadian Life and Health Association of Canada), and the securities sector (through the Investment Dealers Association of Canada, the Mutual Funds Dealers Association, and the Investment Funds Institute of Canada). The NFSO provides Canadian consumers and small businesses access to dispute-resolution services regarding their dealings with financial institutions.

The creation of the NFSO has been viewed as a first step toward building a national regulatory system and eliminating the overlaps among the many federal, provincial, and territorial departments and agencies that currently regulate the different industries of the Canadian financial services sector.

Implications for the Canadian Banking Industry

A bank-holding-company structure (as an alternative to the "bank-as-parent" structure), new ownership rules, expanded permitted investments, expanded access to the payments and clearance system, and a transparent merger review policy offer new opportunities for strategic alliances and joint ventures that have the potential to reshape the Canadian financial services marketplace. These developments, together with new information technologies, make possible new financial products and services and a more vibrant and dynamic market for financial services.

As we have seen, the 1991 federal financial reforms have stimulated consolidation of the Canadian banking industry. The financial consolidation process has further been sped up by the 2001 legislation, because the way is now open to mergers and acquisitions as well as strategic alliances, partnerships, and joint ventures. As already noted, bank financial groups will become not only larger but also increasingly complex organizations, engaging in a full gamut of financial activities.

SUMMARY

1. The history of banking in Canada has left us with a small number of banks chartered by the federal government. Multiple agencies regulate chartered banks: the Office of the Superintendent of Financial Institutions (OSFI), the Bank of Canada, and the Canada Deposit Insurance Corporation (CDIC).
2. The Big Six (the Royal Bank of Canada, Canadian Imperial Bank of Commerce, Bank of Montreal, Scotiabank, TD Canada Trust, and the National Bank of Canada) together with the Desjardins institutions dominate the deposit-taking industry in Canada.
3. In the United States there is a dual banking system, with commercial banks chartered by the states and the federal government. Restrictive state branching regulations that prohibited branching across state lines led to a large number of small commercial banks in the United States. The large number of commercial banks in the United States reflects the past *lack* of competition, not the presence of vigorous competition.
4. A change in the economic environment will stimulate financial institutions to search for financial innovations: Changes in demand conditions, especially an increase

in interest-rate risk; changes in supply conditions, especially improvements in information technology; and the desire to avoid costly regulations have been major driving forces behind financial innovation. Financial innovation has caused banks to suffer declines in cost advantages in acquiring funds and in income advantages on their assets. The resulting squeeze has hurt profitability in banks' traditional line of business and has led to a decline in traditional banking.

5. The regulation and structure of the near banks (trust and loan companies and credit unions and *caisses populaires*) closely parallel the regulation and structure of the chartered banks. Federally incorporated near banks are regulated and supervised by OSFI. They must also register in all the provinces in which they do business and must conform to the regulations of those provinces.
6. With the rapid growth of world trade since 1960, international banking has grown dramatically. Canadian banks engage in international banking activities by opening branches abroad and owning controlling interests in foreign banks. Foreign banks operate in Canada by owning a subsidiary Canadian bank or by operating branches or agency offices in Canada.
7. Until 1981, foreign banks were not allowed to operate in Canada. Today, there are 55 foreign bank subsidiaries and branches operating as Schedule II and III banks in Canada. They have the same powers as the domestic banks but differ in the ownership structure permitted. That is, all Schedule I banks must be widely held, whereas Schedule II and III banks can be closely held if small.
8. The 2001 Bank Act reform introduced a bank-holding-company structure, new ownership rules, expanded access to the payments and clearance system, and new opportunities for strategic alliances and joint ventures. These changes are reshaping the financial services marketplace in Canada by making it easier to introduce new financial products and services and increasing the competitive environment in the industry.

MyLab Economics Make the grade with MyLab Economics: Select end-of-chapter questions can be found on MyLab Economics. You can practise them as often as you want, and most feature step-by-step guided instructions to help you find the right answer.

QUESTIONS

1. Describe how the 2001 Bank Act reform attempted to introduce more competition into Canada's financial services marketplace.
2. Which regulatory agency has the primary responsibility for supervising the following categories of financial institutions?
   a. Chartered banks
   b. Trust and loan companies
   c. Credit unions and *caisses populaires*
3. "The commercial banking industry in Canada is less competitive than the commercial banking industry in the United States because in Canada only a few large banks dominate the industry, while in the United States there are around 4500 commercial banks." Is this statement true, false, or uncertain? Explain your answer.
4. How did new technology cause banks' traditional lending activities to decline in balance-sheet importance?
5. Contrast the activities of a Schedule I bank, a Schedule II bank, a trust company, and a credit union.
6. Explain how sovereign loans in the 1970s and early 1980s caused problems for the Big Six banks.
7. Explain how the early development of chartered banks in Canada differed from the development of commercial banks in the United States.
8. What incentives do Canadian regulatory agencies have to encourage the establishment of foreign banks in Canada?
9. Explain how securitization can be used to change illiquid assets into liquid assets.
10. If the bank at which you keep your chequing account is owned by foreigners, should you worry that your deposits are less safe than if the bank was owned by Canadians?
11. What are the essential differences between chartered banks, trust and loan companies, and credit unions and *caisses populaires*?

12. Why have banks been losing cost advantages in acquiring funds in recent years?

13. "If inflation had not risen in the 1960s and 1970s, the banking industry might be healthier today." Is this statement true, false, or uncertain? Explain your answer.

14. Why have banks been losing income advantages on their assets in recent years?

15. "The invention of the computer is the major factor behind the decline of the banking industry." Is this statement true, false, or uncertain? Explain your answer.

16. Why are there only 2 U.S. banks among the 10 largest banks in the world?

17. What factors explain the rapid growth of international banking?

18. Why has the number of bank holding companies in the United States dramatically increased?

DATA ANALYSIS PROBLEMS

1. Go to the Canada Deposit Insurance Corporation website and access the CDIC deposit insurance calculator.

Use the calculator to find out if your savings are protected by CDIC in the event that your financial institution fails.

2. Go to the CANSIM database and find data on the level of money market investment fund assets (series V37245). Download the data into a spreadsheet.

a. When did assets start entering money market mutual funds? What was the total worth of assets in money market mutual funds at the end of 1977?

b. For each decade period, calculate the total percentage change in assets from the beginning of the decade to the end of the decade: 1980: M1–1990:M1; 1990:M1–2000:M1; 2000: M1–2010:M1; 2010: M1–2020:M1. For each decade period, divide the total percentage change by 10 to get the average yearly percentage increase. Which decade had the largest average yearly growth in money market investment funds?

c. Calculate the growth rate from the most recent month of data available to the same month a year prior. How does this growth rate compare to the highest average yearly growth rate for the decades from part (b)?

3. **Real-Time Data Analysis** Go to the St. Louis Federal Reserve FRED database and find data on the 30-year, fixed-rate average mortgage rate (MORTGAGE30US) and the 5/1-year adjustable-rate mortgage (MORTGAGE5US).

a. What are the mortgage rates reported for the most recent week of data available?

b. If the principal payment for a given month were $2 000, then what would be the interest payment per month (using simple interest) for each of the mortgage types, using the most recent week of data?

c. Over a one-year period, how much would the difference in interest payments between the two mortgage types amount to?

11 Financial Crises

Learning Objectives

11.1 Define the term *financial crisis*.

11.2 Identify the key features of the three stages of a financial crisis.

11.3 Describe the causes and consequences of the global financial crisis of 2007–2009.

11.4 Describe Canada's response to the global financial crisis of 2007–2009

11.5 Summarize the changes to financial regulation that occurred in response to the global financial crisis of 2007–2009.

11.6 Identify the gaps in current financial regulation and how those gaps may be addressed with future regulatory changes.

Preview

Financial crises are major disruptions in financial markets characterized by sharp declines in asset prices and firm failures. Beginning in August 2007, defaults in the subprime mortgage market (by borrowers with weak credit records) sent a shudder through the financial markets, leading to the worst financial crisis since the Great Depression. In congressional testimony, Alan Greenspan, former chairman of the Fed, described the financial crisis as a "once-in-a-century credit tsunami." Wall Street firms and commercial banks suffered hundreds of billions of dollars of losses. Households and businesses found they had to pay higher rates on their borrowings—and it was much harder to get credit. All over the world, stock markets crashed, with the U.S market falling by over 50% from its peak. Many financial firms, including commercial banks, investment banks, and insurance companies, went belly up. A recession began in December 2007, and by the fall of 2008, the economy was in a tailspin. The recession, which ended in June 2009, was the most severe since World War II and is now known as the "Great Recession."

Why did this global financial crisis occur? Why have financial crises been so prevalent throughout Canadian history as well as in so many other countries, and what insights do they provide on the financial crisis from 2007 to 2009? Why are financial crises almost always followed by severe contractions in economic activity, as occurred during the Great Recession? We will examine these questions in this chapter by developing a framework with which to understand the dynamics of financial crises. Building on Chapter 8, we make use of agency theory, the economic analysis of the effects of asymmetric information (adverse selection and moral hazard) on financial markets, to see why financial crises occur and why they have such devastating effects on the economy. We then apply our analysis to explain the course of events that led to a number of past financial crises, including the global financial crisis from 2007 to 2009.[1]

11.1 WHAT IS A FINANCIAL CRISIS?

LO 11.1 Define the term *financial crisis*.

In Chapter 8, we saw that a well-functioning financial system solves asymmetric information problems (moral hazard and adverse selection) so that capital is allocated to its most productive uses. These asymmetric information problems, which act as a barrier

[1]An additional chapter, "Financial Crises in Emerging Market Economies," available online, extends the analysis to financial crises in *emerging market economies*, economies in an early stage of market development that have opened up to the flow of goods, services, and capital from the rest of the world.

to efficient allocation of capital, are often described by economists as **financial frictions**. When financial frictions increase, financial markets are less capable of channelling funds efficiently from savers to households and firms with productive investment opportunities, with the result that economic activity declines. A **financial crisis** occurs when information flows in financial markets experience a particularly large disruption, with the result that financial frictions increase sharply and financial markets stop functioning. Then economic activity collapses.

11.2 DYNAMICS OF FINANCIAL CRISES

LO 11.2 Identify the key features of the three stages of a financial crisis.

As earth-shaking and headline-grabbing as the 2007–2009 financial crisis was, it was only one of a number of financial crises that have hit industrialized countries like Canada and the United States over the years. These experiences have helped economists uncover insights into present-day economic turmoil.

Financial crises in advanced economies have progressed in two and sometimes three stages. To understand how these crises have unfolded, refer to Figure 11-1, which traces the stages and sequence of events in financial crises in advanced economies.

Stage One: Initial Phase

Financial crises can begin in three ways: credit boom and bust, an asset-price boom and bust, or a general increase in uncertainty caused by failures of major financial institutions.

Credit Boom and Bust The seeds of a financial crisis are often sown when an economy introduces new types of loans or other financial products, known as **financial innovation**, or when countries engage in **financial liberalization**, the elimination of restrictions on financial markets and institutions. In the long run, financial liberalization promotes financial development and encourages a well-run financial system that allocates capital efficiently. However, financial liberalization has a dark side: In the short run, it can prompt financial institutions to go on a lending spree, called a **credit boom**. Unfortunately, lenders may not have the expertise, or the incentives, to manage risk appropriately in these new lines of business. Even with proper management, credit booms eventually outstrip the ability of institutions—and government regulators—to screen and monitor credit risks, leading to overly risky lending.

As we learned in Chapter 9, government safety nets, such as deposit insurance, weaken market discipline and increase the moral hazard incentive for banks to take on greater risk than they otherwise would. Because lender-savers know that government-guaranteed insurance protects them from losses, they will supply even undisciplined banks with funds. Without proper monitoring, risk taking grows unchecked.

Eventually, losses on loans begin to mount, and the value of the loans (on the asset side of the balance sheet) falls relative to liabilities, thereby driving down the net worth (capital) of banks and other financial institutions. With less capital, these financial institutions cut back on their lending to borrower-spenders, a process called **deleveraging**. Furthermore, with less capital, banks and other financial institutions become riskier, causing lender-savers and other potential lenders to these institutions to pull out their funds. Fewer funds mean fewer loans to fund productive investments and a credit freeze: The lending boom turns into a lending crash.

Mini-lecture

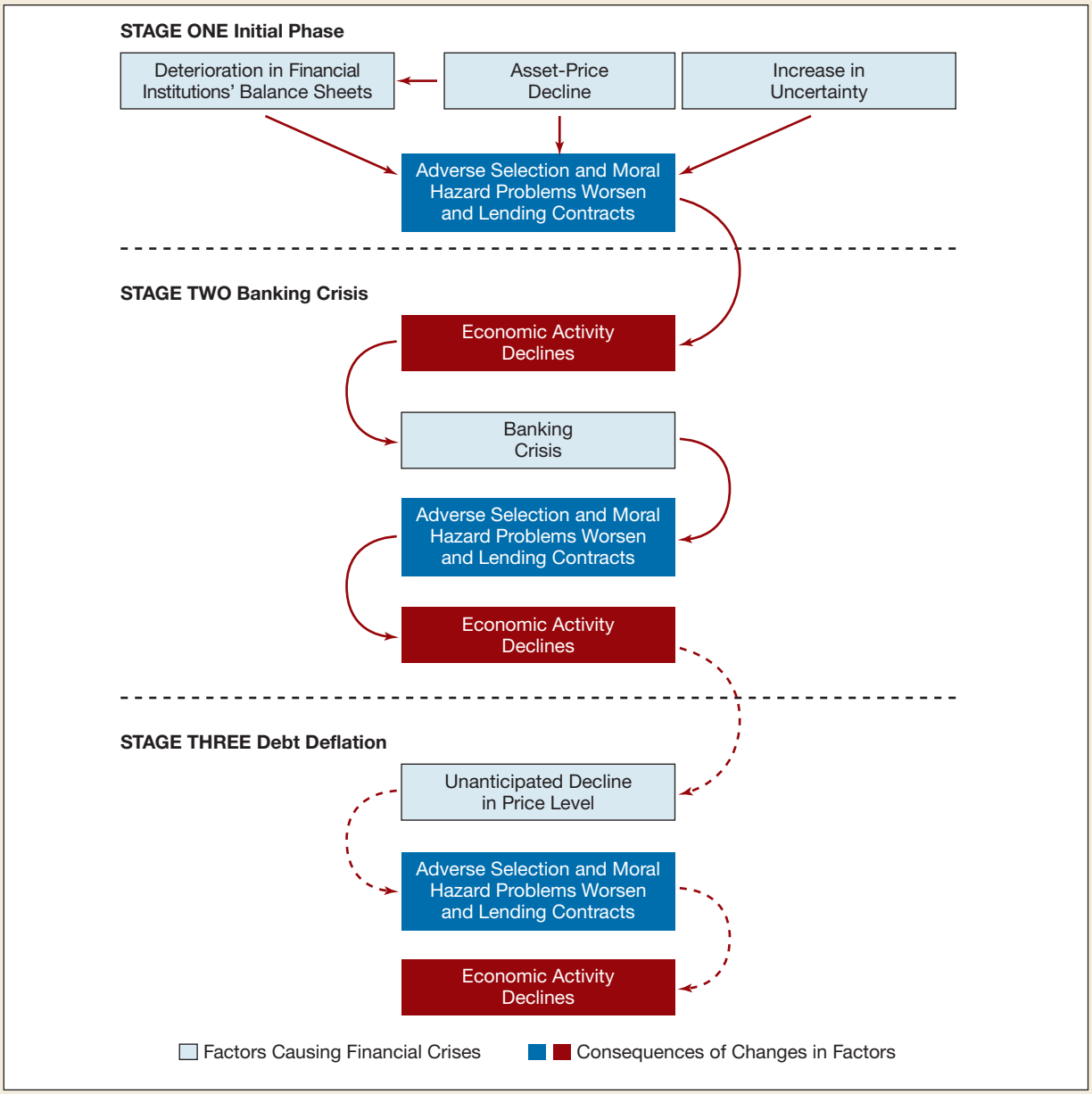

STAGE ONE Initial Phase
Deterioration in Financial Institutions' Balance Sheets
Asset-Price Decline
Increase in Uncertainty
Adverse Selection and Moral Hazard Problems Worsen and Lending Contracts
STAGE TWO Banking Crisis
Economic Activity Declines
Banking Crisis
Adverse Selection and Moral Hazard Problems Worsen and Lending Contracts
Economic Activity Declines
STAGE THREE Debt Deflation
Unanticipated Decline in Price Level
Adverse Selection and Moral Hazard Problems Worsen and Lending Contracts
Economic Activity Declines
Factors Causing Financial Crises
Consequences of Changes in Factors

FIGURE 11-1 Sequence of Events in Financial Crises in Advanced Economies

The solid arrows trace the sequence of events during a typical financial crisis; the dotted arrows show the additional set of events that occurs if the crisis develops into a debt deflation. The sections separated by the dashed horizontal lines show the different stages of a financial crisis.

When financial institutions stop collecting information and making loans, financial frictions rise, limiting the financial system's ability to address the asymmetric information problems of adverse selection and moral hazard (as shown in the arrow pointing from the first factor, "Deterioration in Financial Institutions' Balance Sheets," in the top

row of Figure 11-1). As loans become scarce, borrower-spenders are no longer able to fund their productive investment opportunities, so they decrease their spending, causing economic activity to contract.

Asset-Price Boom and Bust Prices of assets such as equity shares and real estate can be driven by investor psychology (dubbed "irrational exuberance" by Alan Greenspan when he was chairman of the U.S. Federal Reserve) well above their **fundamental economic values**, that is, their values based on realistic expectations of the assets' future income streams. The rise of asset prices above their fundamental economic values is an **asset-price bubble**. Examples of asset-price bubbles are the tech stock market bubble of the late 1990s and the housing price bubble from 2002 to 2006 that we will discuss later in this chapter. Asset-price bubbles are often also driven by credit booms, in which the large increase in credit is used to fund purchases of assets, thereby driving up their price.

When the bubble bursts and asset prices realign with fundamental economic values, stock and real estate prices tumble, companies see their net worth (the difference between their assets and their liabilities) decline, and the value of collateral these companies can pledge drops. Now these companies have less at stake because they have less "skin in the game," and so they are more likely to make risky investments because they have less to lose, the problem of moral hazard. As a result, financial institutions tighten lending standards for these borrower-spenders and lending contracts (as shown by the downward arrow pointing from the second factor, "Asset-Price Decline," in the top row of Figure 11-1).

The asset-price bust also causes a decline in the value of financial institutions' assets, thereby causing a decline in the institutions' net worth and hence a deterioration in their balance sheets (shown by the arrow from the second factor to the first factor in the top row of Figure 11-1), which causes them to deleverage, steepening the decline in economic activity.

Increase in Uncertainty Canadian financial crises have usually began in periods of high uncertainty, such as just after the start of a recession, a crash in the stock market, or the failure of a major financial institution. The failure of major financial institutions (such as Bear Stearns, Lehman Brothers, and AIG in 2008 in the United States) is a particularly important source of heightened uncertainty that features prominently in financial crises. Examples in Canadian history are the Bank of Upper Canada in 1866, the Home Bank in 1923, and the bank panic of 1879. With information hard to come by in a period of high uncertainty, financial frictions increase, reducing lending and economic activity (as shown by the arrow pointing from the last factor, "Increases in Uncertainty," in the top row of Figure 11-1).

Stage Two: Banking Crisis

Deteriorating balance sheets and tougher business conditions lead some financial institutions into insolvency, which happens when their net worth becomes negative. Unable to pay off depositors or other creditors, some banks go out of business. If severe enough, these factors can lead to a bank panic in which multiple banks fail simultaneously. The source of the contagion is asymmetric information. In a panic, depositors, fearing for the safety of their deposits (in the absence of or with limited amounts of deposit insurance) and not knowing the quality of banks' loan portfolios,

withdraw their deposits to the point that the banks fail. Uncertainty about the health of the banking system in general can lead to runs on banks, both good and bad, which forces banks to sell off assets quickly to raise the necessary funds. These **fire sales** of assets may cause their prices to decline so much that more banks become insolvent, and the resulting contagion can then lead to multiple bank failures and a full-fledged bank panic.

With fewer banks operating, information about the creditworthiness of borrower-spenders disappears. Increasingly severe adverse selection and moral hazard problems in financial markets deepen the financial crisis, causing declines in asset prices and the failure of firms throughout the economy that lack funds for productive investment opportunities. Figure 11-1 represents this progression in the stage two portion. Bank panics were a feature of all Canadian financial crises during the nineteenth and twentieth centuries, occurring in 1866, 1879, 1923, and 1930–1933. Bank panics were also a feature of all U.S. financial crises until World War II, occurring every twenty years or so in 1819, 1837, 1857, 1873, 1884, 1893, 1907, and 1930–1933. (The 1933 establishment of federal deposit insurance, which protects depositors from losses, has prevented subsequent bank panics in the United States.)

Eventually, public and private authorities shut down insolvent firms and sell them off or liquidate them. Uncertainty in financial markets declines, the stock market recovers, and balance sheets improve. Financial frictions diminish and the financial crisis subsides. With the financial markets able to operate well again, the stage is set for an economic recovery.

Stage Three: Debt Deflation

If, however, the economic downturn leads to a sharp decline in the price level, the recovery process can be short-circuited. In stage three of Figure 11-1, **debt deflation** occurs when a substantial unanticipated decline in the price level sets in, leading to a further deterioration in firms' net worth because of the increased burden of indebtedness.

In economies with moderate inflation, which characterizes most advanced countries, many debt contracts with fixed interest rates are typically of fairly long maturity, 10 years or more. Because debt payments are contractually fixed in nominal terms, an unanticipated decline in the price level raises the value of borrowing firms' and households' liabilities in real terms (increases the burden of the debt) but does not raise the real value of their assets. The borrowers' net worth in real terms (the difference between assets and liabilities in real terms) thus declines.

To better understand how this decline in net worth occurs, consider what happens if a firm in 2022 has assets of $100 million (in 2022 dollars) and $90 million of long-term liabilities, so it has $10 million in net worth (the difference between the values of assets and liabilities). If the price level falls by 10% in 2023, the real value of the liabilities would rise to $99 million in 2022 dollars, while the real value of the assets would remain unchanged at $100 million. The result would be that real net worth in 2022 dollars would fall from $10 million to $1 million ($100 million minus $99 million).

The substantial decline in the real net worth of borrowers caused by a sharp drop in the price level creates an increase in adverse selection and moral hazard problems for lenders. Lending and economic activity decline for a long time. The most significant financial crisis that displayed debt deflation was the Great Depression, the worst economic contraction in Canadian history.

APPLICATION

The Mother of All Financial Crises: The Great Depression in the United States

With our framework for understanding financial crises in place, we are prepared to analyze how a financial crisis unfolded during the Great Depression and how it led to the worst economic downturn in U.S. history.

Stock Market Crash

In 1928 and 1929, U.S. stock prices doubled. Federal Reserve officials viewed the stock market boom as caused by excessive speculation. To curb it, they pursued a tightening of monetary policy to raise interest rates in an effort to limit the rise in stock prices. The Fed got more than it bargained for when the stock market crashed in October 1929, falling by 40% by the end of 1929, as shown in Figure 11-2.

Bank Panics

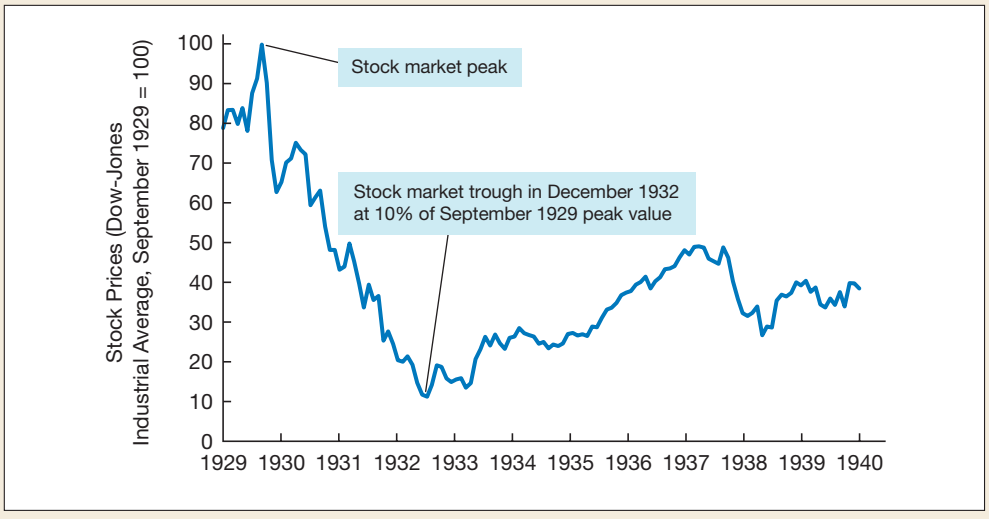
By the middle of 1930, stocks had recovered almost half of their losses and credit market conditions had stabilized. What might have been a normal recession turned into something far worse, however, when severe droughts in the Midwest led to a sharp decline in agricultural production, with the result that farmers could not pay back their bank loans. The resulting defaults on farm mortgages led to large loan losses on bank balance sheets in agricultural regions. The general weakness of the economy, and of the banks in agricultural regions in particular, prompted substantial withdrawals from banks, building to a full-fledged panic in November and December 1930, with the stock market falling sharply. For more than two years, the Fed sat idly by through one bank panic after another, the most severe spate of panics in U.S. history. After what would be the era's final panic in March 1933, President Franklin Delano Roosevelt declared a bank

FIGURE 11-2

Stock Prices During the Great Depression Period

Stock prices crashed in 1929, falling by 40% by the end of 1929, and then continued to fall to only 10% of their peak value by 1932.

Source: Federal Reserve Bank of St. Louis FRED database: fred.stlouisfed.org/series/M1109BUSM293NNBR.



holiday, a temporary closing of all banks. "The only thing we have to fear is fear itself," Roosevelt told the nation. The damage was done, however, and more than one-third of U.S. commercial banks had failed.

Continuing Decline in Stock Prices

Stock prices kept falling. By mid-1932, stocks had declined to 10% of their value at the 1929 peak (as shown in Figure 11-2), and the increase in uncertainty from the unsettled business conditions created by the economic contraction worsened adverse selection and moral hazard problems in financial markets. With a greatly reduced number of financial intermediaries still in business, adverse selection and moral hazard problems intensified even further. Financial markets struggled to channel funds to borrower-spenders with productive investment opportunities. The amount of outstanding commercial loans fell by half from 1929 to 1933, and investment spending collapsed, declining by 90% from its 1929 level.

A manifestation of the rise in financial frictions is that lenders began charging businesses much higher interest rates to protect themselves from credit losses. The resulting rise in the **credit spread**—the difference between the interest rate on loans to households and businesses and the interest rate on completely safe assets that are sure to be paid back, such as U.S. Treasury securities—is shown in Figure 11-3, which displays the difference between interest rates on corporate bonds with a Baa (medium-quality) credit rating and rates on similar-maturity Treasury bonds. (Note the credit spread is closely related to the risk premium discussed in Chapter 6.)

Debt Deflation

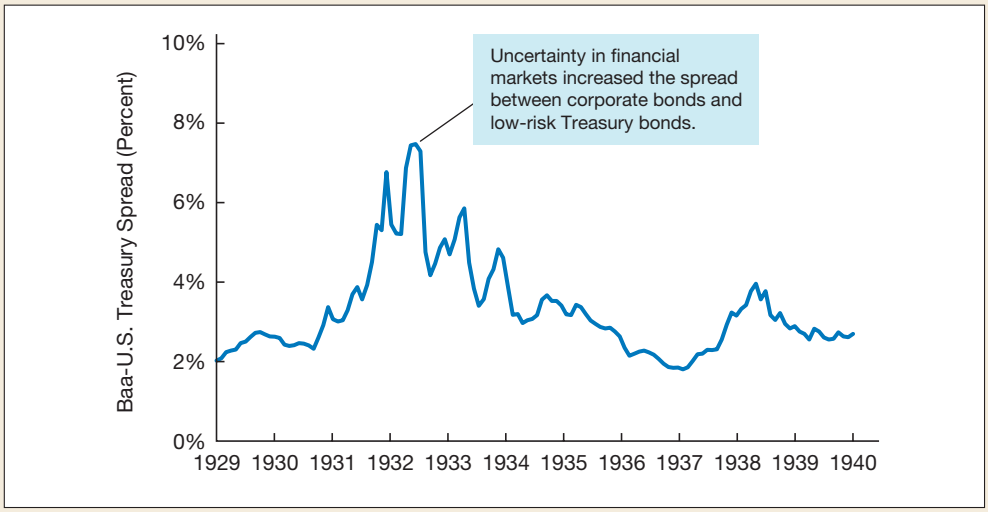
The ongoing deflation that accompanied declining economic activity eventually led to a 25% decline in the price level. This deflation short-circuited the normal recovery process that occurs in most recessions. The huge decline in prices triggered a debt deflation in which real net worth fell because of the increased burden of indebtedness borne by firms and households. The decline in net worth and the resulting increase in adverse selection

FIGURE 11-3

Credit Spreads During the Great Depression in the United States

Credit spreads (the difference between rates on Baa corporate bonds and U.S. Treasury bonds) rose sharply during the Great Depression.

Source: Banking and Monetary Statistics 1914–1941 (Board of Governors of the Federal Reserve, 1943).


Baa-U.S. Treasury Spread (Percent)
10%
8%
6%
4%
2%
0%
1929 1930 1931 1932 1933 1934 1935 1936 1937 1938 1939 1940
Uncertainty in financial markets increased the spread between corporate bonds and low-risk Treasury bonds.

and moral hazard problems in the credit markets led to a prolonged economic contraction in which unemployment rose to 25% of the labour force. The financial crisis of the Great Depression was the worst ever experienced in the United States, which explains why the economic contraction was also the most severe ever experienced by that nation.

International Dimensions

Although the Great Depression started in the United States, it was not just a U.S. phenomenon. Bank panics in the United States also spread to the rest of the world, and the contraction of the U.S. economy sharply decreased the demand for foreign goods. The worldwide depression caused great hardship, with millions upon millions of people out of work, and the resulting discontent led to the rise of fascism and World War II. The consequences of the Great Depression financial crisis were disastrous. ◆

11.3 THE GLOBAL FINANCIAL CRISIS OF 2007–2009

LO 11.3 Describe the causes and consequences of the global financial crisis of 2007–2009.

For many years, most economists thought that financial crises of the type experienced during the Great Depression were a thing of the past for advanced countries like Canada and the United States. Unfortunately, the financial crisis that engulfed the world in 2007–2009 proved them wrong.

Causes of the 2007–2009 Financial Crisis

We begin our look at the 2007–2009 financial crisis by examining three central factors: financial innovation in mortgage markets, agency problems in mortgage markets, and the role of asymmetric information in the credit-rating process.

Financial Innovation in the Mortgage Markets As we saw in Chapter 10, starting in the early 2000s, advances in information technology made it easier to securitize subprime mortgages, leading to an explosion in subprime mortgage-backed securities. Financial innovation didn't stop there. Financial engineering, the development of new, sophisticated financial instruments, led to **structured credit products** that paid out income streams from a collection of underlying assets, designed to have particular risk characteristics that appealed to investors with differing preferences. The most notorious of these products were collateralized debt obligations (CDOs) (discussed in the FYI box "Collateralized Debt Obligations [CDOs]").

Agency Problems in the Mortgage Markets The mortgage brokers who originated the mortgage loans often did not make a strong effort to evaluate whether a borrower could pay off the mortgage, since they planned to quickly sell (distribute) the loans to investors in the form of mortgage-backed securities. This originate-to-distribute business model was exposed to the principal–agent problem of the type discussed in Chapter 8, in which the mortgage brokers acted as agents for investors (the principals) but did not have the investors' best interests at heart. Once the mortgage broker earns their fee, why should the broker care if the borrower makes good on the payment? The more volume the broker originates, the more money the broker makes.

FYI Collateralized Debt Obligations (CDOs)

The creation of a collateralized debt obligation involves a corporate entity called a *special purpose vehicle* (*SPV*), which buys a collection of assets such as corporate bonds and loans, commercial real estate bonds, and mortgage-backed securities. The SPV then separates the payment streams (cash flows) from these assets into a number of buckets that are referred to as *tranches*. The highest-rated tranches, called *super senior tranches*, are the ones that are paid off first and so have the least risk. The super senior CDO is a bond that pays out these cash flows to investors, and because it has the least risk, it also has the lowest interest rate. The next bucket of cash flows, known as the *senior tranche*, is paid out next; the senior CDO has a little more risk and pays a higher interest rate. The next tranche of payment streams, the *mezzanine tranche* of the CDO, is paid out after the super senior and senior tranches, and so it bears more risk and has an even higher interest rate. The lowest tranche of the CDO is the *equity tranche*; this is the first set of cash flows that are not paid out if the underlying assets go into default and stop making payments. This tranche has the highest risk and is often not traded.

If all of this sounds complicated, it is. Tranches also included CDO^2s and CDO^3s that sliced and diced risk even further, paying out the cash flows from CDOs to CDO^2s and from CDO^2s to CDO^3s. Although financial engineering carries the potential benefit of creating products and services that match investors' risk appetites, it also has a dark side. Structured products like CDOs, CDO^2s, and CDO^3s can get so complicated that it becomes hard to value the cash flows of the underlying assets for a security or to determine who actually owns these assets. Indeed, in October 2007, Ben Bernanke, then chairman of the U.S. Federal Reserve, joked that he "would like to know what those damn things are worth." In other words, the increased complexity of structured products can actually reduce the amount of information in financial markets, thereby worsening asymmetric information in the financial system and increasing the severity of adverse selection and moral hazard problems.

Not surprisingly, adverse selection became a major problem. Risk-loving real estate investors lined up to obtain loans to acquire houses that would be very profitable if housing prices went up, knowing they could "walk away" if housing prices went down. The principal–agent problem also created incentives for mortgage brokers to encourage households to take on mortgages they could not afford or to commit fraud by falsifying information on borrowers' mortgage applications in order to qualify them for mortgages. Compounding this problem was lax regulation of originators, who were not required to disclose information to borrowers that would have helped them assess whether they could afford the loans.

The agency problems went even deeper. Commercial and investment banks, which were earning large fees by underwriting mortgage-backed securities and structured credit products like CDOs, also had weak incentives to make sure that the ultimate holders of the securities would be paid off. **Financial derivatives**, financial instruments whose payoffs are linked to (i.e., derived from) previously issued securities, also were an important source of excessive risk taking. Large fees from writing a type of financial insurance contract called a **credit default swap**, a financial derivative that provides payments to holders of bonds if they default, also drove units of insurance companies like AIG to write hundreds of billions of dollars' worth of these risky contracts.

Asymmetric Information and Credit-Rating Agencies Credit-rating agencies, who rate the quality of debt securities in terms of the probability of default, were another contributor to asymmetric information in financial markets. The rating agencies advised clients on how to structure complex financial instruments, like CDOs, while at the same time they were rating these identical products. The rating agencies were thus subject to conflicts of interest because the large fees they earned from advising clients on how to structure products that they themselves were rating meant that they did not have sufficient incentives to make sure their ratings were accurate. The result was wildly inflated ratings that enabled the sale of complex financial products that were far riskier than investors recognized.

Effects of the 2007–2009 Financial Crisis

Consumers and businesses alike suffered as a result of the 2007–2009 financial crisis. The impact of the crisis was most evident in five key areas: the U.S. residential housing market, financial institutions' balance sheets, the shadow banking system, global financial markets, and the headline-grabbing failures of major firms in the financial industry.

Residential Housing Prices: Boom and Bust The subprime mortgage market took off after the recession ended in the United States in 2001. By 2007, it had become over a trillion-dollar market. The development of the subprime mortgage market was encouraged by economists and politicians alike because it led to a "democratization of credit" and helped raise U.S. homeownership rates to the highest levels in history. The asset-price boom in housing (see Figure 11-4), which took off after the 2000–2001

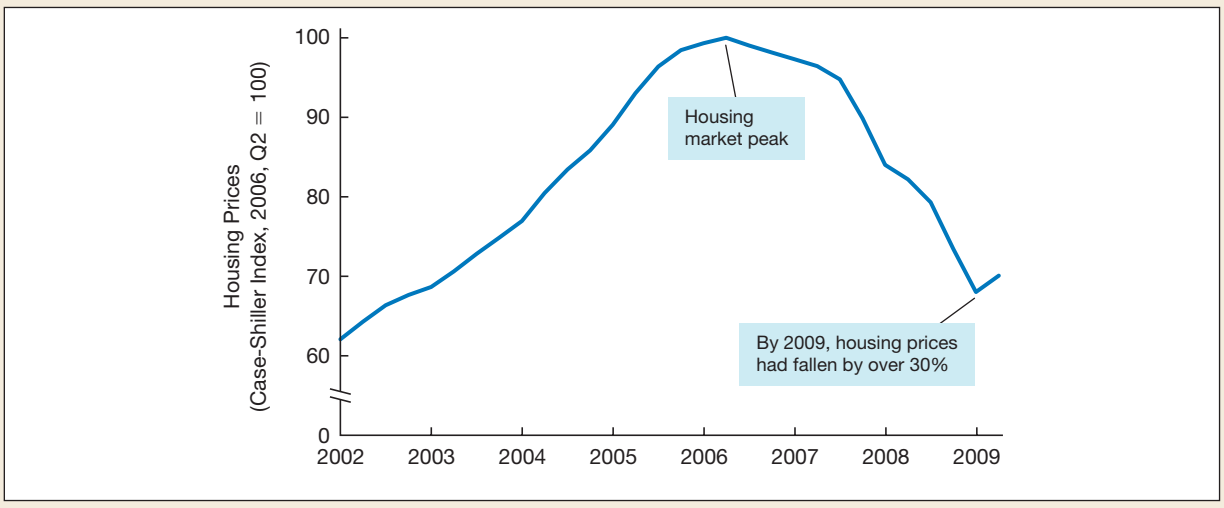


FIGURE 11-4 Housing Prices in the United States and the Financial Crisis of 2007–2009

Housing prices boomed from 2002 to 2006, fuelling the market for subprime mortgages and forming an asset-price bubble. Housing prices began declining in 2006, falling by more than 30% subsequently, which led to defaults by subprime mortgage holders.

Source: Case-Shiller U.S. National Composite House Price Index from Federal Reserve Bank of St. Louis FRED database: fred.stlouisfed.org/series/SPCS20RSA.

FYI Was the Fed to Blame for the Housing Price Bubble?

Some economists—most prominently, John Taylor of Stanford University—have argued that the low interest rate policy of the U.S. Federal Reserve in the 2003–2006 period caused the housing price bubble.* Taylor argues that the low federal funds rate led to low mortgage rates that stimulated housing demand and encouraged the issuance of subprime mortgages, both of which led to rising housing prices and a bubble.

In a speech given in January 2010, then-Federal Reserve Chairman Ben Bernanke countered this argument.† He concluded that monetary policy was not to blame for the housing price bubble. First, he said, it is not at all clear that the federal funds rate was too low during the 2003–2006 period. Rather, the culprits were the proliferation of new mortgage products that lowered mortgage payments, a relaxation of lending standards that brought more buyers into the housing market, and capital inflows from countries such as China and India. Bernanke's speech was very controversial, and the debate over whether monetary policy was to blame for the housing price bubble continues to this day.

*John Taylor, "Housing and Monetary Policy," in Federal Reserve Bank of Kansas City, *Housing, Housing Finance and Monetary Policy* (Kansas City: Federal Reserve Bank of Kansas City, 2007), 463–476.

†Ben S. Bernanke, "Monetary Policy and the Housing Bubble," speech given at the annual meeting of the American Economic Association, Atlanta, Georgia, January 3, 2010; www.federalreserve.gov/newsevents/speech/bernanke20100103a.htm.

recession was over, also helped stimulate the growth of the subprime mortgage market. High housing prices meant that subprime borrowers could refinance their houses with even larger loans when their homes appreciated in value. With housing prices rising, subprime borrowers were also unlikely to default because they could always sell their house to pay off the loan, making investors happy because the securities backed by cash flows from subprime mortgages had high returns. The growth of the subprime mortgage market, in turn, increased the demand for houses and so fuelled the boom in housing prices, resulting in a housing price bubble.

Further stimulus for the inflated housing market came from low interest rates on residential mortgages, which were the result of several different forces. First were the huge capital inflows into the United States from countries like China and India. Second was congressional legislation that encouraged Fannie Mae and Freddie Mac to purchase trillions of dollars of mortgage-backed securities.[2] Third was Federal Reserve monetary policy that made it easy to lower interest rates. The low cost of financing for housing purchases that resulted from these forces further stimulated the demand for housing, pushing up housing prices. (A highly controversial issue is whether the Federal Reserve was to blame for the housing price bubble, and this is discussed in the FYI box.)

As housing prices rose and profitability for mortgage originators and lenders grew higher, the underwriting standards for subprime mortgages fell lower and lower. High-risk borrowers were able to obtain mortgages, and the amount of the mortgage relative to the value of the house, the loan-to-value ratio, rose. Borrowers were often able to get piggyback, second, and third mortgages on top of their original 80% loan-to-value mortgage so that they had to put almost no money down. When asset prices rise too far out of line

[2]For a discussion of the government's role in encouraging the housing boom, which led to a subsequent bust in the housing market, see Thomas Sowell, *The Housing Boom and Bust*, rev. ed. (New York: Basic Books, 2010).

with fundamentals—in the case of housing, the cost of purchasing a home relative to the cost of renting it, or the cost of houses relative to households' median income—they must come down. Eventually, the housing price bubble burst. With housing prices falling after their peak in 2006 (see Figure 11-4), the rot in the financial system began to reveal itself. The decline in housing prices led many subprime borrowers to find that their mortgages were "underwater"—that is, the value of the house was below the amount of the mortgage. When this happened, struggling homeowners had tremendous incentives to walk away from their homes and just send the keys back to the lender. Defaults on mortgages shot up sharply, eventually leading to foreclosures on millions of mortgages.

Deterioration of Financial Institutions' Balance Sheets The decline in U.S. housing prices led to rising defaults on mortgages. As a result, the values of mortgage-backed securities and CDOs collapsed, leaving banks and other financial institutions holding those securities with a lower value of assets and thus a lower net worth. With weakened balance sheets, these banks and other financial institutions began to deleverage, selling off assets and restricting the availability of credit to both households and businesses. With no one else able to step in to collect information and make loans, the reduction in bank lending meant that financial frictions increased in financial markets.

Run on the Shadow Banking System The sharp decline in the values of mortgages and other financial assets triggered a run on the shadow banking system, composed of hedge funds, investment banks, and other nondepository financial firms, which are not as tightly regulated as banks. Funds from shadow banks flowed through the financial system and for many years supported the issuance of low-interest-rate mortgages and auto loans.

These securities were funded primarily by **repurchase agreements (repos)**, short-term borrowing that, in effect, uses assets like mortgage-backed securities as collateral. Rising concern about the quality of a financial institution's balance sheet led lenders to require larger amounts of collateral, known as **haircuts**. For example, if a borrower took out a $100 million loan in a repo agreement, the borrower might have to post $105 million of mortgage-backed securities as collateral, for a haircut of 5%.

Rising defaults on mortgages caused the values of mortgage-backed securities to fall, which then led to a rise in haircuts. At the start of the crisis, haircuts were close to zero, but eventually they rose to nearly 50%.[3] The result was that financial institutions could borrow only half as much with the same amount of collateral. Thus, to raise funds, financial institutions had to engage in fire sales and sell off their assets very rapidly. Because selling assets quickly requires lowering their price, the fire sales led to a further decline in financial institutions' asset values. This decline lowered the value of collateral further, raising haircuts and thereby forcing financial institutions to scramble even more for liquidity. The result was similar to the run on the banking system that occurred during the Great Depression, causing massive deleveraging that resulted in a restriction of lending and a decline in economic activity.

The decline in asset prices in the stock market (which fell by over 50% from October 2007 to March 2009, as shown in Figure 11-5) and the more than 30% drop in residential house prices (shown in Figure 11-4), along with the fire sales resulting from the run on the shadow banking system, weakened both firms' and households' balance sheets. This worsening of financial frictions manifested itself in widening credit spreads, causing higher costs of credit for households and businesses and tighter

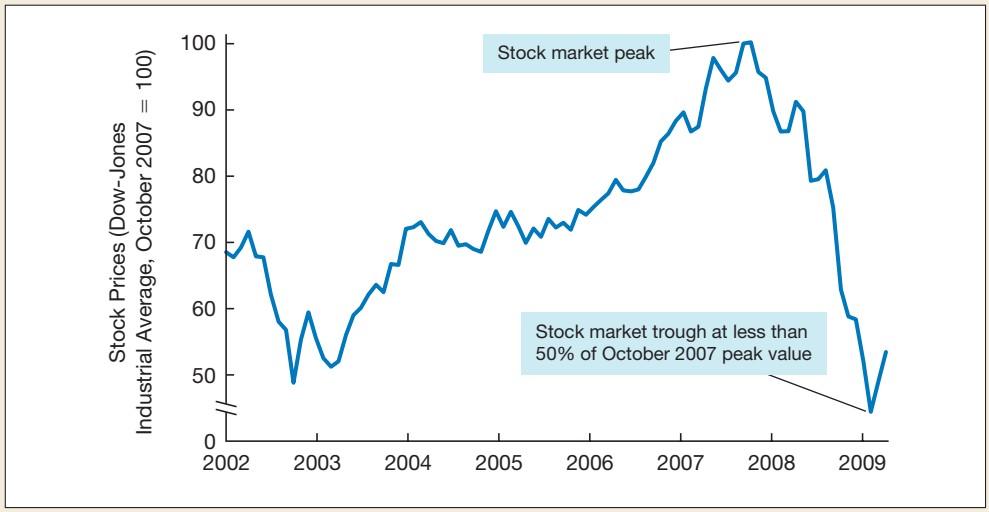
[3]See Gary Gorton and Andrew Metrick, "Securitized Banking and the Run on Repo," *Journal of Financial Economics* 104, no. 3 (2012): 425–451.

FIGURE 11-5

Stock Prices in the United States and the Financial Crisis of 2007–2009

Stock prices fell by 50% from October 2007 to March 2009.

Source: Federal Reserve Bank of St. Louis FRED database, fred.stlouisfed.org/series/DJIA.



lending standards. The resulting decline in lending meant that both consumption expenditure and investment fell, causing the economy to contract.[4]

Global Financial Markets Although the problem originated in the United States, the wake-up call for the financial crisis came from Europe, a sign of how extensive the globalization of financial markets had become. After Fitch and Standard & Poor's announced ratings downgrades on mortgage-backed securities and CDOs totalling more than US $10 billion, on August 7, 2007, a French investment house, BNP Paribas, suspended redemption of shares held in some of its money market funds, which had sustained large losses. The run on the shadow banking system began, only to become worse and worse over time. Despite huge injections of liquidity into the financial system by the European Central Bank and the U.S. Federal Reserve, banks began to horde cash and were unwilling to lend to each other. The drying up of credit led to the first major bank failure in the United Kingdom in over 100 years when Northern Rock, which had relied on short-term borrowing in the repo market rather than deposits for its funding, collapsed in September 2007. A string of other European financial institutions then failed as well. Particularly hard hit were countries like Greece, Ireland, Portugal, and Spain. The resulting crisis in markets for government-issued (sovereign) debt in Europe is described in the Global box "The European Sovereign Debt Crisis."

Failure of High-Profile Firms The impact of the financial crisis on firms' balance sheets forced major players in the financial markets to take drastic action. In March 2008, Bear Stearns, the fifth-largest investment bank in the United States, which had invested heavily in subprime-related securities, had a run on its repo funding and was forced to sell itself to J.P. Morgan for less than one-tenth of its worth just a year earlier. To broker the deal, the Federal Reserve had to take over US$30 billion of Bear Stearns's

[4]This period was also characterized by a substantial supply shock that occurred when oil and other commodity prices rose sharply until the summer of 2008 but then fell precipitously thereafter. We discuss the aggregate demand and supply analysis of the impact of this supply shock in Chapter 23.

Global The European Sovereign Debt Crisis

The global financial crisis of 2007–2009 led not only to a worldwide recession but also to a sovereign debt crisis that still threatens to destabilize Europe today. Up until 2007, all of the countries that had adopted the euro found their interest rates converging to very low levels, but with the onset of the global financial crisis, several of these countries were hit very hard by the contraction in economic activity, which reduced tax revenues at the same time that government bailouts of failed financial institutions required additional government outlays. The resulting surge in budget deficits then led to fears that the governments of these hard-hit countries would default on their debt. The result was a surge in interest rates that threatened to spiral out of control.*

Greece was the first domino to fall in Europe. In September 2009, with an economy weakened by reduced tax revenues and increased spending demands, the Greek government was projecting a budget deficit for the year of 6% and a debt-to-GDP ratio near 100%. However, when a new government was elected in October, it revealed that the budget situation was far worse than anyone had imagined because the previous government had provided misleading numbers about both the budget deficit, which was at least double the 6% number, and the amount of government debt, which was 10 percentage points higher than previously reported. Despite austerity measures aimed at dramatically cutting government spending and raising taxes, interest rates on Greek debt soared, eventually rising to nearly 40%, and the debt-to-GDP ratio climbed to 160% of GDP in 2012. Even with bailouts from other European countries and liquidity support from the European Central Bank, Greece was forced to write down the value of its debt held in private hands by more than half, and the country was subject to civil unrest, with massive strikes and the resignation of the prime minister.

The sovereign debt crisis spread from Greece to Ireland, Portugal, Spain, and Italy. The governments of these countries were forced to embrace austerity measures to shore up their public finances while interest rates climbed to double-digit levels. Only with a speech in July 2012 by Mario Draghi, the president of the European Central Bank, in which he stated that the ECB was ready to do "whatever it takes" to save the euro, did the markets begin to calm down. Nonetheless, despite a sharp decline in interest rates in these countries, they experienced severe recessions, with unemployment rates rising to double-digit levels and Spain's unemployment rate exceeding 25%. The stresses that the European sovereign debt crisis produced for the euro zone, the countries that adopted the euro, at that time raised doubts about whether the euro would survive.

*For a discussion of the dynamics of sovereign debt crises and case studies of the European debt crisis, see David Greenlaw, James D. Hamilton, Frederic S. Mishkin, and Peter Hooper, "Crunch Time: Fiscal Crises and the Role of Monetary Policy," *U.S. Monetary Policy Forum* (Chicago: Chicago Booth Initiative on Global Markets, 2013).

hard-to-value assets. In July, Fannie Mae and Freddie Mac, the two privately owned, government-sponsored enterprises that together insured over US$5 trillion of mortgages or mortgage-backed assets, were propped up by the U.S. Treasury and the Federal Reserve after suffering substantial losses from their holdings of subprime securities. In early September 2008, Fannie Mae and Freddy Mac were put into conservatorship (in effect, run by the government).

On Monday, September 15, 2008, after suffering losses in the subprime market, Lehman Brothers, the fourth-largest investment bank by asset size with over US$600 billion in assets and 25 000 employees, filed for bankruptcy, making it the largest bankruptcy filing in U.S. history. The day before, Merrill Lynch, the third-largest investment bank, which had also suffered large losses on its holdings of subprime securities, announced its sale to Bank of America for a price 60% below its value a year earlier.

On Tuesday, September 16, AIG, an insurance giant with assets of over US$1 trillion, suffered an extreme liquidity crisis when its credit rating was downgraded. It had written over US$400 billion of insurance contracts (credit default swaps) that had to make payouts on possible losses from subprime mortgage securities. The Federal Reserve then stepped in with a US$85 billion loan to keep AIG afloat (with total government loans later increased to US$173 billion).

Height of the 2007–2009 Financial Crisis

The financial crisis reached its peak in September 2008 after the U.S. House of Representatives, fearing the wrath of constituents who were angry about the Wall Street bailout, voted down a US$700 billion dollar bailout package proposed by the Bush administration. The Emergency Economic Stabilization Act was finally passed nearly a week later. The stock market crash accelerated, with the week beginning October 6, 2008, showing the worst weekly decline in U.S. history. Credit spreads went through the roof over the next three weeks, with the spread between Baa corporate bonds (just above investment grade) and U.S. Treasury bonds going to over 5.5 percentage points (550 basis points), as illustrated by Figure 11-6.

The impaired financial markets and surging interest rates faced by borrower-spenders led to sharp declines in consumer spending and investment. Real GDP declined sharply, falling at a −1.3% annual rate in the third quarter of 2008 and then at a −5.4% and −6.4% annual rate in the next two quarters. The unemployment rate shot up, going over the 10% level in late 2009. The recession that started in December 2007 became the worst economic contraction in the United States since World War II and, as a result, is now referred to as the "Great Recession."

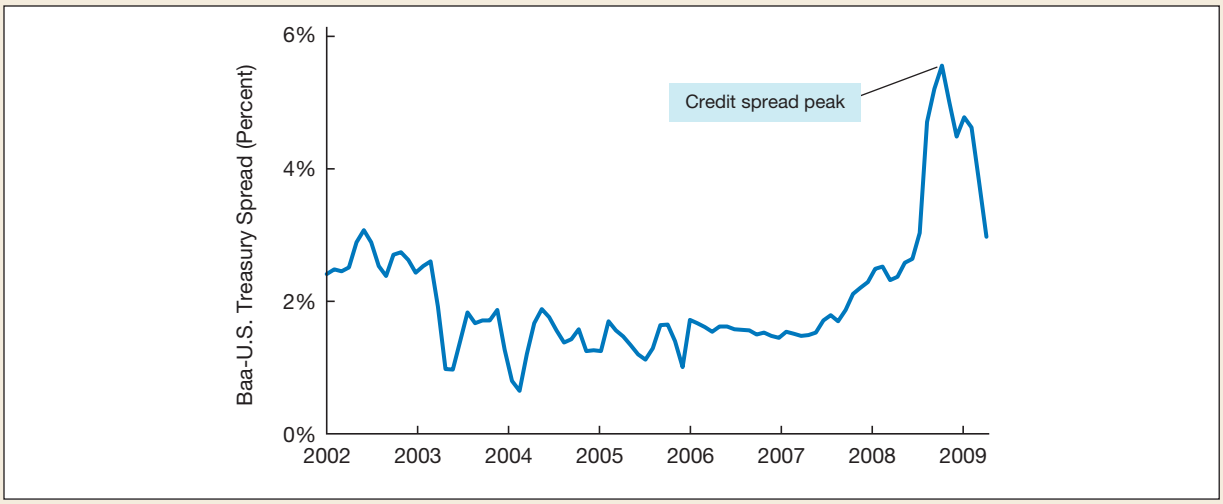


FIGURE 11-6 Credit Spreads and the 2007–2009 Financial Crisis in the United States

Credit spreads (the difference between rates on Baa corporate bonds and U.S. Treasury bonds) rose by more than four percentage points (400 basis points) during the crisis. Debate over the bailout package and the stock market crash caused credit spreads to peak in December 2008.

Source: Federal Reserve Bank of St. Louis FRED database: fred.stlouisfed.org/series/BAA10Y.

Government Intervention and the Recovery

Although the recession produced by the global financial crisis was very severe, the economic contraction was far smaller in magnitude than during the Great Depression because of massive interventions by governments to prop up financial markets and stimulate the economy.

As we will see in Chapter 16, the Federal Reserve took extraordinary actions to contain the crisis, actions involving both monetary policy to stimulate the economy and liquidity provision to support orderly functioning of financial markets. In addition, the U.S. government engaged in massive bailouts, with over US$150 billion of loans to AIG and the **Troubled Asset Relief Program (TARP)**, the most important provision of the Bush administration's Emergency Economic Stabilization Act passed in October 2008, which authorized the Treasury to spend $700 billion purchasing subprime mortgage assets from troubled financial institutions or to inject capital into these institutions—the route actually followed. In addition, the act raised the federal deposit insurance limit temporarily from $100 000 to $250 000 in order to limit withdrawals from banks. Shortly thereafter, the FDIC put in place a guarantee for certain debt newly issued by banks, and the Treasury guaranteed for a year money market mutual fund shares at par value. Similarly, European governments conducted massive bailouts, in excess of US$10 trillion, to prop up their banking systems (see the Global box "Worldwide Government Bailouts During the 2007–2009 Financial Crisis").

Fiscal policy aimed at stimulating the economy was another key piece of the U.S. government's response to the crisis. In February 2008, Congress passed the Bush administration's Economic Stimulus Act of 2008, whereby the government gave out one-time tax rebates totalling US$78 billion by sending $600 cheques to individual taxpayers. Shortly after coming to office, the Obama administration proposed the American Recovery and Reinvestment Act of 2009, a much bigger, $787 billion fiscal stimulus package passed by Congress that to this day is highly controversial. This stimulus package is discussed more extensively in Part 6 of the book.

Global Worldwide Government Bailouts During the 2007–2009 Financial Crisis

The spreading bank failures in Europe in the fall of 2008 led to massive bailouts of financial institutions: The Netherlands, Belgium, and Luxembourg injected US$16 billion to prop up Fortis, a major European bank; the Netherlands injected US$13 billion into ING, a banking and insurance giant; Germany provided a US$50 billion rescue package for Hypo Real Estate Holdings; and Iceland took over its three largest banks after its banking system collapsed. Ireland's government guaranteed all the deposits of its commercial banks as well as interbank lending, as did Greece. Spain implemented a bailout package similar to that of the United States, in which it bought up to 50 billion euros (US$70 billion) of assets in its banks in order to encourage them to lend. The U.K. Treasury set up a bailout plan similar to that of the U.S. Treasury's plan with a price tag of 400 billion pounds (US$699 billion). It guaranteed 250 billion pounds of bank liabilities, added 100 billion pounds to a facility that swaps these assets for government bonds, and allowed the U.K. government to buy up to 50 billion pounds of equity stakes in British banks. Bailout plans to the tune of over US$100 billion in South Korea, US$200 billion in Sweden, US$400 billion in France, and US$500 billion in Germany, all of which guaranteed the debt of their banks as well as the injection of capital into them, then followed. Both the scale of these bailout packages and the degree of international coordination were unprecedented.

With the government bailouts, the Fed's extraordinary actions, and fiscal stimulus, a bull market in stocks got under way starting in March 2009 (see Figure 11-5), and credit spreads began to fall (Figure 11-6). With the recovery of financial markets, the economy also started to recover, but unfortunately the pace of recovery has been slow.

APPLICATION Could the Coronavirus Pandemic Have Led to a Financial Crisis?

The coronavirus pandemic in 2020 had the potential to trigger a financial crisis as serious as the 2007–2009 global financial crisis. With the start of the lockdown of economies around the world in March 2020, stock markets crashed, falling by more than a third, unemployment skyrocketed, and many otherwise healthy firms faced the prospect of being unable to pay their bills or pay back their loans. The framework we laid out in our discussion of Figure 11-1 can be used to analyze how the coronavirus pandemic provided the seeds for another financial crisis, 12 years after the previous one, and why it didn't become the next financial crisis.

All the factors that potentially lead to a financial crisis, shown in the first row of Figure 11-1, came into play when the pandemic became severe in March 2020. The lockdowns dealt a serious blow to the income of both businesses and households, making it more likely that they would be unable to pay back their loans. A severe deterioration in financial institutions' balance sheets thus became a real possibility, which could have led to severe restrictions on lending. The stock market crash resulted in more than a 35% decline in stock prices, and the sharp decline in income of a multitude of businesses produced a sharp drop in the net worth of firms, which would increase both adverse selection and moral hazard. The high degree of uncertainty about the spread of the virus and how long it would disrupt the economy increased the asymmetry of information, making it harder to assess credit risks.

The seeds of a financial crisis were then planted, and indeed, credit spreads such as the Baa-Treasury spread in the United States doubled, shooting up from 2 percentage points in February to a peak of 4.3 percentage points on March 23, 2020. Although the coronavirus pandemic had the potential to trigger a full-fledged financial crisis in Canada and other countries, this did not occur. The reason was the massive response by both governments and central banks. Both governments and central banks around the world reacted with unprecedented speed once the World Health Organization announced that the spread of the coronavirus could now be classified as a pandemic on March 11, 2020. In March, the Bank of Canada not only slashed its policy rate (overnight interest rate) to 0.25%, which is what it considers to be its effective lower bound, but also embarked on large-scale programs to stabilize the financial markets (see Chapter 16). The Canadian government also introduced a large number of programs to support individuals and businesses, such as the Canada Emergency Response Benefit, the Canada Emergency Wage Subsidy, the Canada Emergency Rent Subsidy, among others. This massive (almost $250 billion) stimulus package provided loans and grants to small businesses and large corporations, aid to provincial governments, and increases in employment insurance.

The coronavirus pandemic had the potential to unleash another financial crisis, with disastrous effects on the economy. However, the combination of central bank and government policies helped shore up businesses, and credit spreads began to decline. As of this writing, the impact of the coronavirus pandemic on the economy is still uncertain, but the likelihood of a financial crisis has decreased substantially. ◆

11.4 CANADA AND THE 2007–2009 FINANCIAL CRISIS

LO 11.4 Describe Canada's response to the global financial crisis of 2007–2009.

In Canada, an early symptom of the U.S. subprime crisis was the freezing of the asset-backed commercial paper (ABCP) market in August 2007. It caused a sharp decrease in liquidity in short-term Canadian credit markets, including the overnight interbank market and the overnight repurchase market. In order to restore confidence and liquidity in the ABCP market, and prevent significant write-offs for banks that would have reduced their capital and raised concerns about their solvency, market participants reached an agreement, known as the Montreal Accord, to restructure the frozen ABCP market and minimize the costs to the Canadian taxpayer. (See the FYI box "Canada's Asset-Backed Commercial Paper Crisis" for more details.)

FYI Canada's Asset-Backed Commercial Paper Crisis

As we noted in Chapter 2, commercial paper is short-term, unsecured debt issued by corporations, typically for financing accounts receivable and inventories. Asset-backed commercial paper, like traditional commercial paper, is also a short-term security (with a maturity that is typically less than nine months). It is issued by conduits (that is, bankruptcy-remote Special Purpose Vehicles), but instead of being an unsecured promissory note, it is backed by physical assets such as mortgages, trade receivables, credit card receivables, automobile loans and leases, and other types of assets. Because of this backing, the quality of the ABCP depends on the underlying securities and thus the ABCP could be very risky. For example, if there are negative developments in the underlying markets, the risk of the ABCP will increase and investors will switch out of the ABCP and into safer money market instruments such as traditional commercial paper, bankers' acceptances, and Treasury bills. As a result, conduits will not be able to roll over their ABCP and will face a liquidity crunch.

The ABCP market in Canada expanded very rapidly from the 1990s to 2007. As at July 31, 2007, the size of the Canadian ABCP market was $115 billion (equal to about 32.5% of the Canadian money market), of which $80 billion was bank-sponsored and $35 billion was non-bank-sponsored. At the same time, the ABCP market in the United States was approximately US$1.2 trillion, equal to about 50% of the U.S. commercial paper market. Bank-sponsored ABCP conduits are invested in "plain vanilla" assets such as residential mortgages and credit card receivables. Non-bank-sponsored ABCP conduits are invested in structured finance assets such as CDOs and subprime mortgages. Canadian banks are involved in the distribution of ABCP and also provide liquidity back-stop facilities to non-bank-sponsored ABCP conduits under the so-called general market disruption clause.

In August 2007, investors in the Canadian ABCP market declined to roll over maturing notes because of concerns about exposure to the U.S. subprime mortgage sector in the underlying assets. As a result, the market was divided into those ABCP conduits that could honour their obligations (bank-sponsored) and those that could not honour their obligations (non-bank-sponsored).

In the case of bank-sponsored ABCP, the Big Six banks took back onto their balance sheets significant amounts of their own sponsored ABCP. Moreover, valuation and fair-value issues led to significant writedowns in the fourth quarter of 2007 and the first quarter of 2008. In the case of non-bank-sponsored ABCP, a number of conduits faced significant liquidity shortages, seeking liquidity funding support from their liquidity providers (banks). However, major liquidity providers denied requests for liquidity support, arguing that the "general market disruption" clause was not met. They interpreted the clause to mean that the majority of the conduits in the entire Canadian ABCP market would need to be unable to roll over before a liquidity provider had to step in. As a result, the Canadian non-bank-sponsored ABCP

market froze, and investors, including Québec's huge pension fund, *Caisse de dépôt et placement du Québec,* Alberta's ATB Financial, the National Bank of Canada, and about 2000 individual small investors, had their cash frozen in non-bank-sponsored ABCP.

At the time, the Bank of Canada indicated that it would not accept ABCP as collateral for loans to banks and that a solution from participants in the ABCP market was deemed to be appropriate. As a result, major market participants, including non-bank-sponsored ABCP conduits, institutional investors, liquidity providers (ABN AMRO Group, HSBC, Deutsche Bank, Merrill Lynch, Barclays Capital), and the affected trusts, reached an agreement on August 16, 2007, known as the Montreal Accord. Under the Montreal Accord, investors agreed to a standstill period (initially 60 days, to October 16, 2007, extended three times to February 22, 2008), with the objective of restructuring the frozen ABCP into long-term floating rate notes with maturities matching the maturities of the underlying assets in the conduits. However, the agreement had to be renegotiated again 16 months later, in December 2008, with the participation of the federal government and three provinces (Ontario, Québec, and Alberta), where publicly owned institutions held over $30 billion of frozen non-bank-sponsored ABCP. Small investors (holding less than $1 million of the paper) were fully paid out in cash by the brokerage firms that had sold them the paper, whereas the larger investors were given bonds.

The ABCP crisis in Canada was managed well, although the final cost of the restructuring was in excess of $200 million in fees, with the big investors footing the bill.

Subprime Mortgages in Canada

The subprime mortgage crisis in the United States dominated news headlines, but there is also a Canadian version of subprime mortgages. High-risk, long-term (40-year), zero-down mortgages proliferated in Canada in 2007 and 2008 after the Conservative government opened up the Canadian mortgage market to big U.S. players in its first budget in May 2006. This created the Canadian version of subprime mortgages. How did Canadian regulators allow the entry into Canada of U.S.-style subprime mortgages?

The Canadian mortgage insurance market is the second largest and one of the most lucrative mortgage insurance markets in the world. For over 50 years, since revisions in the National Housing Act and the Bank Act in 1954 allowed chartered banks to make insured mortgage loans, the business of mortgage insurance was mainly the domain of the Canada Mortgage and Housing Corporation (CMHC). CMHC insured the mortgages of those home buyers who could not make a 25% down payment, charging them an insurance premium, with the federal government backing the CMHC's insurance policies with a 100% federal guarantee.

For many years, U.S. insurance companies had conducted orchestrated lobbying of Canada's Department of Finance and Office of the Superintendent of Financial Institutions. In May 2006, the government allowed big U.S. players such as AIG to enter the Canadian mortgage insurance market and also committed to guarantee their business up to $200 billion. In February 2006, in an attempt to protect its business in anticipation of the entry of the big U.S. insurers, CMHC announced that it would insure 30-year mortgages. Two weeks later, Genworth Mortgage Insurance Co., the Canadian mortgage subsidiary of the U.S. conglomerate General Electric and the only other (private) mortgage insurer in Canada at that time (with an estimated market share of about 30%), announced that it would insure 35-year mortgages. By October 2006, AIG, CMHC, and Genworth were competing in the Canadian mortgage insurance market by insuring 40-year mortgages.

In response to the subprime meltdown in the United States, the Canadian government banned subprime mortgages in Canada in the summer of 2008. Although there

are currently no publicly available figures, it is estimated that more than half of the total new mortgages approved by Canadian financial institutions during this period (worth over $50 billion) were risky 40-year mortgages and that 10% of these mortgages (worth over $10 billion) were taken out with zero money down.

In the aftermath of the global financial crisis, Canadian regulatory authorities tightened several regulations for real estate lending to ensure that Canadians could not take on bigger mortgages than they could afford if interest rates were to rise. The most significant regulatory changes are as follows: the maximum amortization period for a mortgage is now 25 years; refinancing a home will be allowed only up to 80% of its appraised value; and homebuyers will have to demonstrate their housing costs are no more than 35% of their gross household income. In addition, Canada's banking regulator, the Office of the Superintendent of Financial Institutions (OSFI), has placed a loan-to-value limit of 65% (of a home's purchase price) on borrowing against home equity. Finally, in 2018, OSFI introduced a new mortgage qualification rule requiring lenders to stress test all uninsured mortgage loans (those with a down payment of 20% or more). As of June 1, 2021, the qualifying rate on uninsured mortgages is either 200 basis points above the negotiated contract rate or 5.25%, whichever is higher.

Why Was Canada Exempt from the Financial Crisis?

Canadian banks also had their problems during the turbulent financial conditions of the global financial crisis. Their shares fell by almost 50%, and some of them experienced huge losses in derivatives trading; for example, CIBC lost $2.1 billion in derivatives trading in 2008. However, while governments in the United States and Europe worked on full-scale banking bailouts and rescue packages (in the trillions of dollars), the Canadian government did not have to bail out any banks. Canada was the only G7 country to experience a mild recession (even milder than the 1980s and early 1990s recessions), with no bank failures and bailouts of insolvent firms. This is why, in the aftermath of the global financial crisis, Canada has been ranked first among more than 140 countries in terms of banking stability by the World Economic Forum.

One reason that Canada's banks have fared better than banks in the United States and other countries is that they are more tightly regulated. For example, depending on its charter, a bank in the United States can be regulated by the Federal Reserve, the Federal Deposit Insurance Corporation, the Office of the Comptroller of the Currency, or state regulators. By contrast, Canada's top banking regulator, the Office of the Superintendent of Financial Institutions oversees all financial firms, including banks, insurance companies, mortgage lenders, and pension funds, among others. Moreover, although Canadian securities markets are regulated by the country's provincial and territorial governments, their regulations are highly harmonized. In addition, bank regulations in Canada are revised every five years in order to help regulation adapt to emerging risks and innovation.

Another reason is that Canada's big banks have been more conservative in their lending and acquisition practices compared to banks in the United States and major banks around the world. In fact, the Office of the Superintendent of Financial Institutions has been more conservative than banking regulators in the United States and Europe. For example, at the beginning of the global financial crisis, Canada's banks had higher capital requirements, greater leverage restrictions, and fewer off-balance-sheet activities than their global peers. As a result, they had stronger reserves to cushion potential losses. Although this conservative regulatory regime enabled Canadian banks to withstand the financial crisis better than banks in other countries, it has been argued

that it makes the Canadian banking sector less competitive because of the lower leverage and a lower rate of return on capital than in other jurisdictions.

Moreover, the activities of Canada's banks are well diversified and are not limited to traditional retail banking. In particular, the federal government's decision in the late 1980s to allow banks to acquire investment brokers on Bay Street and to engage in the mutual fund and insurance businesses created a more diversified financial services marketplace for Canada's banks. In addition, these arm's-length institutions are subject to the same strict rules and regulations as the banks, unlike investment dealers in the United States that had been subject to very relaxed and minimal regulation from the Securities and Exchange Commission. As a result, shadow banking activity in Canada today accounts for only about 40% of the size of Canada's economy, compared with 95% in the United States.

Finally, Canada's structural and regulatory differences manifest themselves very well in the structure of the Canadian mortgage market. Unlike banks in the United States that sold the bulk of their mortgages to investors, banks in Canada tend to hold a large proportion of their mortgages on their balance sheets; before the global financial crisis, for example, less than a third of Canadian mortgages were securitized, compared with two-thirds in the United States. This practice gives Canadian banks an incentive to make sure that their mortgage loans are good loans. In addition, law in Canada allows banks to go after other assets when a consumer walks away from a mortgage, thereby making it difficult for consumers to do so.

Overall, in the aftermath of the global economic meltdown, Canada's banking system has been viewed as the soundest in the world. In fact, many countries around the world are now considering Canadian-style reforms of their financial markets.

11.5 RESPONSE OF FINANCIAL REGULATION

LO 11.5 Summarize the changes to financial regulation that occurred in response to the global financial crisis of 2007–2009.

Given the cost to the economy of the 2007–2009 financial crisis, the size of the bailouts, and the nationalization of so many financial institutions, the system of financial regulation in the United States is currently undergoing dramatic changes.

Macroprudential Versus Microprudential Supervision

Before the global financial crisis, regulatory authorities engaged in **microprudential supervision**, which focuses on the safety and soundness of *individual* financial institutions. Microprudential supervision looks at each individual institution separately and assesses the riskiness of its activities and whether it complies with disclosure requirements. Most importantly, microprudential supervision checks whether a particular institution satisfies capital *ratios* and, if it does not, either engages in prompt corrective action to force the institution to raise its capital ratios or the supervisor closes it down, along the lines we have discussed in Chapter 9.

A focus on microprudential supervision is not enough to prevent financial crises. The run on the shadow banking system illustrates how the problems of one financial institution can harm other financial institutions that are otherwise healthy. When the troubled financial institution is forced to engage in fire sales and sell off assets in order to meet target capital ratios or haircut requirements, this leads to a decline in asset values. The decline in

asset values then causes other institutions to engage in fire sales, leading to a rapid deleveraging process and a systemic crisis. In situations like this, even institutions that have high capital ratios and would normally be healthy may find themselves in trouble.

The global financial crisis has therefore made it clear that there is a need for **macroprudential supervision**, which focuses on the safety and soundness of the financial system *in the aggregate*. Rather than focusing on the safety and soundness of individual institutions, macroprudential supervision seeks to mitigate systemwide fire sales and deleveraging by assessing the overall capacity of the financial system to avoid them. In addition, because many institutions that were well capitalized faced liquidity shortages and found that their access to short-term funding was cut off, macroprudential supervision focuses not only on capital adequacy as a whole but also on whether the financial system has sufficient liquidity.

Macroprudential policies can take several forms. The run-up to the global financial crisis included a so-called **leverage cycle**, in which a feedback loop resulted from a boom in issuing credit, which led to higher asset prices, which resulted in higher capital buffers at financial institutions, which supported further lending in the context of unchanging capital requirements, which then raised asset prices further, and so on; in the bust, the opposite occurs, with the value of the capital dropping precipitously, leading to a cut in lending. To short-circuit this leverage cycle, macroprudential policies make capital requirements countercyclical; that is, they are adjusted upward during a boom and downward during a bust.

In addition, during the upward swing in the leverage cycle, macroprudential policies might involve forcing financial institutions to tighten credit standards or even direct limits on the growth of credit. In the downward swing, macroprudential supervision might be needed to force the banking system as a whole to raise an aggregate amount of new capital so that banks will not curtail lending in order to reduce the level of their assets and raise capital ratios. To ensure that financial institutions have enough liquidity, macroprudential policies could require that financial institutions have a sufficiently low *net stable funding ratio*, which is the percentage of the institution's short-term funding in relation to total funding. Macroprudential policies of the type discussed here are being considered as part of the Basel 3 framework but have not yet been completely worked out.

Dodd-Frank Wall Street Reform and Consumer Protection Act of 2010

The global financial crisis raised calls for a new regulatory structure that would make a repeat of the crisis less likely. The result was the Dodd-Frank bill, which was passed in July 2010 after more than a year of discussion. It is the most comprehensive financial reform legislation passed in the United States since the Great Depression. The bill addresses seven different categories of regulation, which are discussed next.

Consumer Protection Dodd-Frank created a new Consumer Financial Protection Bureau that is funded and housed within the Federal Reserve, although it is a completely independent agency. It has the authority to examine and enforce regulations on all businesses with more than $10 billion in assets that are engaged in issuing residential mortgage products as well as on issuers of other financial products marketed to low-income people. The legislation requires lenders to make sure borrowers can repay residential mortgages by requiring verification of their income, credit history, and job status. It also bans payments to brokers for pushing borrowers into higher-priced loans. It allows states to impose stricter consumer protection laws on national banks and gives

state attorneys general the power to enforce certain rules issued by the new bureau. It also permanently increased the level of federal deposit insurance to $250 000.

Annual Stress Tests Dodd-Frank requires that banks with assets of more than $10 billion be subject every year to **stress tests**, a supervisory assessment of whether banks have sufficient bank capital if they are faced with a scenario of bad macroeconomic outcomes, such as a collapse in housing prices or a severe recession. Not only are the results of these stress tests published every year but if a bank "fails" the stress test and is found to have insufficient capital, it also has to restrict the amount of dividends it pays out and come up with a plan to raise new capital to eliminate this deficiency.

Stress tests were first conducted for 19 banks at the height of the financial crisis in the first half of 2009. When the results were announced in May 2009, they were well received by market participants, enabling these banks to raise substantial amounts of capital from private capital markets. The stress tests were a key factor that helped increase the amount of information in the marketplace, thereby reducing asymmetric information and adverse selection and moral hazard problems. Their success led to Dodd-Frank making them a regular exercise on an annual basis.

Resolution Authority Before the new legislation was passed, the FDIC had the authority to seize failing banks and wind them down, but it did not have such resolution authority over the largest financial institutions—those structured as holding companies. Indeed, the U.S. Treasury and the Federal Reserve argued that one reason they were unable to rescue Lehman Brothers, and instead had to let it go into bankruptcy, was that they did not have the legal means with which to take over Lehman and break it up. The Dodd-Frank bill now provides the U.S. government this resolution authority, called the Orderly Liquidation Authority, for financial firms that are deemed systemic—that is, firms that pose a risk to the overall health of the financial system because their failure would cause widespread economic damage. It also gives regulators the right to levy fees on financial institutions with more than $50 billion in assets to recoup any losses.

Limits on Federal Reserve Lending As we have seen, during the financial crisis, the Federal Reserve made loans to individual firms such as Bear Stearns and AIG under an emergency lending authority. Because of concerns that this Fed lending to individual firms entailed a bailout that would encourage firms to take on excessive risk in the future, the Dodd-Frank bill directed the Federal Reserve to limit emergency lending to "broad-based" programs rather than to individual firms and also required the Fed to get approval from the U.S. Treasury Department to make use of its emergency lending authority.

Systemic Risk Regulation The Dodd-Frank bill created a Financial Stability Oversight Council, chaired by the Treasury secretary, that monitors markets for asset-price bubbles and the buildup of systemic risk. In addition, it designates which financial firms are systemically important and so receive the official designation of **systemically important financial institutions (SIFIs)**. These firms are subject to additional regulation by the Federal Reserve, which includes higher capital standards and stricter liquidity requirements, along with the requirement that such firms draw up a "living will," that is, a plan for orderly liquidation if the firm gets into financial difficulties.

Volcker Rule Under the new legislation, banks are limited in the extent of their proprietary trading—that is, trading with their own money—and are allowed to own only a small percentage of hedge and private equity funds. These provisions are named

after Paul Volcker, a former chairman of the Board of Governors of the Federal Reserve, who argued that banks should not be allowed to take large trading risks when they receive the benefits of federal deposit insurance.

Derivatives As discussed in Chapter 10, derivatives such as credit default swaps ended up being "weapons of mass destruction" that helped lead to a financial meltdown when AIG had to be rescued after making overly extensive use of them. To prevent this from happening again, the Dodd-Frank bill requires that many standardized derivative products be traded on exchanges to make their trading more transparent and also requires that they be cleared through clearinghouses to reduce the risk of losses if one counterparty in the derivative transaction goes bankrupt. More customized derivative products are subject to higher capital requirements. Banks are banned from some of their derivative-dealing operations, such as those involving riskier swaps. In addition, the bill imposes capital and margin requirements on firms dealing in derivatives and forces them to disclose more information about their activities.

11.6 TOO BIG TO FAIL AND FUTURE REGULATION

LO 11.6 Identify the gaps in current financial regulation and how those gaps may be addressed with future regulatory changes.

The Dodd-Frank bill leaves out many of the details of future regulation, and there are doubts as to whether it has dealt sufficiently with the too-big-to-fail problem, which, as we have seen, was an important factor that contributed to the global financial crisis. Here we discuss some possible measures to reduce the too-big-to-fail problem and explore several areas in which regulation might be heading in the future.

What Can Be Done About the Too-Big-to-Fail Problem?

Three approaches to solving the too-big-to-fail problem have been actively debated.

Break Up Large, Systemically Important Financial Institutions One way to eliminate the too-big-to-fail problem is to make sure that no financial institution is so large that it can bring down the financial system. Then, regulators will no longer see the need to bail out these institutions if they fail, thereby subjecting them to market discipline. One way to shrink overly large institutions is to reimpose the restrictions that were in place before Glass-Steagall was repealed, thereby forcing these large SIFIs to break up their different activities into smaller, cohesive companies. Alternatively, regulations could specify that no financial institution can have assets over a specified maximum limit, forcing SIFIs to break up into smaller pieces.

Not surprisingly, both of these approaches have been vehemently opposed by the largest financial institutions. Although breaking up SIFIs would eliminate the too-big-to-fail problem, if there are synergies available that might enable large institutions to manage risk better or to provide financial services at a lower cost, then breaking up SIFIs might decrease the efficiency of the financial system rather than increase it.

Higher Capital Requirements Because institutions that are "too big to fail" have incentives to take on excessive risk, another way to reduce their risk taking is to impose higher capital requirements on them. With higher capital, not only will these

institutions have a larger buffer with which to withstand losses if they occur but they will also have more to lose and hence have more "skin in the game," thereby reducing moral hazard and giving them less incentive to take on excessive risk. Another way of describing this approach is to say that higher capital requirements reduce the subsidy to risk taking for institutions that are too big to fail. In addition, because risk taking by SIFIs is far greater during booms, capital requirements could be increased when credit is expanding rapidly and reduced when credit is contracting. Such measures would cause capital requirements to become more countercyclical and could help restrain the boom-bust credit cycle.

The Swiss central bank has been a leader in this kind of approach: It has the highest capital requirements among advanced countries for its largest banks, and it raises them when credit markets become particularly frothy. Legislation has been proposed in the U.S. Congress to double capital requirements for large financial institutions, legislation that these institutions vigorously oppose.

Leave It to Dodd-Frank Another view is that Dodd-Frank has effectively eliminated the too-big-to-fail problem by making it harder for the Federal Reserve to bail out financial institutions, by imposing stricter regulations on SIFIs, and through application of the Volcker rule. Indeed, the authors of the bill have declared that Dodd-Frank will "end too-big-to-fail as we know it." Although the provisions of Dodd-Frank do take away some of the incentives for excessive risk taking by large, systemically important financial institutions, there are doubts that this bill completely removes the too-big-to-fail problem.

Beyond Dodd-Frank: Where Might Regulation Head in the Future?

In February 2017, the new president, Donald Trump, signed an executive order directing the secretary of the Treasury to review the financial regulations in the Dodd-Frank bill. Critics of Dodd-Frank, such as the Trump administration, have argued that many of its regulations have restricted banks from lending, especially to households, making it harder for them to borrow. Supporters of Dodd-Frank argue that it has made the financial system sounder, reducing the likelihood of a financial crisis. Four years later the executive order signed by Trump was revoked by President Joe Biden. Here we examine debates about where regulation should head in the future in several areas.

Consumer Protection Congressional critics of the Consumer Financial Protection Bureau argue that its regulations, such as those requiring lenders to make sure that borrowers have the ability to repay loans and mortgages, have led banks to be overly conservative in extending credit to consumers. Some of these critics would even go beyond restraining the Consumer Financial Protection Bureau and would like to see it abolished, arguing that this would promote more lending to households. Advocates for consumer protection have vowed to fight to retain these regulations and the Consumer Financial Protection Bureau because they argue that not only will these regulations and the Consumer Financial Protection Bureau protect consumers from fraud and deception, but they will also make the financial system safer because fraudulent practices helped promote the subprime lending boom and the global financial crisis.

Resolution Authority There are proposals in the U.S. Congress to abolish the Orderly Liquidation Authority. Advocates for these proposals argue that the Orderly

Liquidation Authority legitimizes federal bailouts of large financial firms and therefore increases the too-big-to-fail problem. Instead, they propose an enhanced bankruptcy procedure for failing financial firms that would be determined by the courts. Defenders of a federal resolution authority argue that although the Orderly Liquidation Authority can be improved, it is still needed because bankruptcy, even an enhanced procedure, is a messy legal process that makes orderly resolutions of failing financial institutions in a crisis extremely difficult. Without an Orderly Liquidation Authority, the failure of a major financial institution is more likely to cause a financial crisis to spin out of control.

Volcker Rule The Volcker rule to limit banks from proprietary trading with their own money, which has the advantage of limiting excessive risk taking by banks, has come under attack for several reasons. First, regulators have found it very difficult to differentiate proprietary trading from banks trading for their customers, so the Volcker rule has resulted in very complex regulations that might be very difficult and costly to enforce. Second, it limits banks from engaging in a profitable activity that could offset possible losses from lending. Third, critics of the Volcker rule argue that by limiting trading, it has reduced liquidity in many financial markets. In 2018, the Federal Reserve loosened some of the restrictions imposed by the Volcker rule.

Derivatives Trading Proposals in the U.S. Congress seek to roll back the Dodd-Frank rules on derivatives trading. Their proponents argue that these rules have hurt the ability of financial institutions in the United States to make these trades, with the result that derivatives trading, particularly in credit default swaps, has moved overseas, making it harder for American financial firms to make profits and reducing the ability of U.S. nonfinancial firms to insure against risk. Supporters of the Dodd-Frank rules on derivatives trading argue that they make derivatives trading more transparent and safer, thereby reducing the likelihood of future financial crises.

Government-Sponsored Enterprises (GSEs) A major gap in the Dodd-Frank bill is that it does not address privately owned, government-sponsored enterprises, such as Fannie Mae and Freddie Mac. During the global financial crisis, both of these firms got into serious financial trouble and had to be taken over by the government, with massive loans and government guarantees needed to prop them up. To prevent this from occurring again, there are proposals to reform the GSEs by four different routes:

1. Fully privatize GSEs by taking away their government sponsorship, thereby removing the implicit backing for their debt.
2. Completely nationalize them by taking away their private status and making them government agencies.
3. Leave them as privately owned GSEs, but strengthen regulations to restrict the amount of risk they take and to impose higher capital standards.
4. Leave them as privately owned GSEs, but force them to shrink dramatically so that they no longer expose the taxpayer to huge losses or pose a systemic risk to the financial system when they fail.

However, supporters of the current status of the GSEs argue against these reforms because they would limit GSE activity, thereby making it harder for households to get credit.

SUMMARY

1. A financial crisis occurs when a particularly large disruption to information flows occurs in financial markets, with the result that financial frictions increase sharply, thereby rendering financial markets incapable of channelling funds to households and firms with productive investment opportunities and causing a sharp contraction in economic activity.
2. Financial crises start in advanced countries like Canada and the United States in two ways: credit booms and busts, or a general increase in uncertainty when major financial institutions fail. The result is a substantial increase in adverse selection and moral hazard problems, which leads to a contraction of lending and a decline in economic activity. The worsening business conditions and deterioration in bank balance sheets then trigger the second stage of the crisis, the simultaneous failure of many banking institutions, a banking crisis. The resulting decrease in the number of banks causes a loss of their information capital, leading to a further decline of lending and a spiralling down of the economy. In some instances, the resulting economic downturn leads to a sharp slide in prices, which increases the real liabilities of firms and households and therefore lowers their net worth, leading to a debt deflation. The further decline in borrowers' net worth worsens adverse selection and moral hazard problems so that lending, investment spending, and aggregate economic activity remain depressed for a long time.
3. The most significant financial crisis in history, which led to the Great Depression, involved several stages: a stock market crash, bank panics, worsening of asymmetric information problems, and finally a debt deflation.
4. The global financial crisis of 2007–2009 was triggered by mismanagement of financial innovations involving subprime residential mortgages and the bursting of a housing price bubble. The crisis spread globally, with substantial deterioration in banks' and other financial institutions' balance sheets, a run on the shadow banking system, and the failure of many high-profile firms.
5. During the 2007–2009 global financial crisis, unlike the United States and Europe, Canada only experienced a mild recession. However, in the aftermath of the crisis, Canadian regulatory authorities tightened several regulations for real estate lending in Canada.
6. Financial regulation has undergone many changes in the wake of the global financial crisis. First is the shift from microprudential supervision, which focuses on the safety and soundness of individual institutions, to macroprudential supervision, which focuses on the safety and soundness of the financial system in the aggregate. Second is the most comprehensive financial reform legislation since the Great Depression. It makes provisions in seven areas: (1) consumer protection, (2) annual stress tests, (3) resolution authority, (4) limits on central bank lending, (5) systemic risk regulation, (6) the Volcker rule in the United States, and (7) derivatives trading.
7. Future regulation needs to address several issues: (1) the too-big-to-fail problem, which can be at least partially addressed either by breaking up large financial institutions or by imposing higher capital requirements on them; (2) many of the provisions of the Dodd-Frank legislation in the United States; and (3) reform of the GSEs to make it less likely that they will require government bailouts in the future.

MyLab Economics Make the grade with MyLab Economics: Select end-of-chapter questions can be found on MyLab Economics. You can practise them as often as you want, and most feature step-by-step guided instructions to help you find the right answer.

QUESTIONS

1. How does the concept of asymmetric information help to define a financial crisis?
2. How can the bursting of an asset-price bubble in the stock market help trigger a financial crisis?
3. How does an unanticipated decline in the price level cause a drop in lending?
4. Define "financial frictions" in your own terms and explain why an increase in financial frictions is a key element in financial crises.

5. How does a deterioration in balance sheets of financial institutions and the simultaneous failures of these institutions cause a decline in economic activity?
6. How does a general increase in uncertainty as a result of the failure of a major financial institution lead to an increase in adverse selection and moral hazard problems?
7. What is a credit spread? Why do credit spreads rise significantly during a financial crisis?
8. Some countries do not advertise that a system of deposit insurance like the CDIC in Canada and the FDIC in the United States exists in their banking system. Explain why some countries would want to do that.
9. Describe the process of "securitization" in your own words. Was this process solely responsible for the Great Recession financial crisis of 2007–2009?
10. Provide one argument in favour of and one against the idea that the Fed was responsible for the U.S. housing price bubble of the mid-2000s.
11. What role does weak financial regulation and supervision play in causing financial crises?
12. Describe two similarities and two differences between the United States' experiences during the Great Depression and the Great Recession financial crisis of 2007–2009.
13. What do you think prevented the financial crisis of 2007–2009 from becoming a depression?
14. What technological innovations led to the development of the subprime mortgage market?
15. Why is the originate-to-distribute business model subject to the principal–agent problem?
16. True, false, or uncertain: Deposit insurance always and everywhere prevents financial crises.
17. How did a decline in housing prices help trigger the subprime financial crisis in the United States that began in 2007?
18. What role did the shadow banking system play in the 2007–2009 financial crisis?
19. Why would haircuts on collateral increase sharply during a financial crisis? How would this lead to fire sales on assets?
20. How did the global financial crisis promote a sovereign debt crisis in Europe?
21. Why is it a good idea for macroprudential policies to require countercyclical capital requirements?
22. How does the process of financial innovation affect the effectiveness of macroprudential regulation?
23. What are the three approaches to limiting the too-big-to-fail problem? Briefly describe the advantages and disadvantages of each approach.
24. What was the Canadian experience during the 2007–2009 global financial crisis?

DATA ANALYSIS PROBLEMS

1. Go to the CANSIM database and find data (over the period from 2000 to 2020) on household final consumption expenditure (series V61989012), stock prices (series V122620), and house prices (series V111955443). Download the data into a spreadsheet and convert the monthly stock price and house price series to quarterly frequency. Make sure the data align correctly with the appropriate dates. For all three series, for each quarter, calculate the annualized growth rate from quarter to quarter. To do this, take the current-period data minus the previous-quarter data, and then divide by the previous quarter data. Multiply by 100 to change each result to a percentage, and multiply by 4 to annualize the data.
    a. For the three series, calculate the growth rates over the first three quarters of 2020. Comment on the relationships among house prices, stock prices, and consumption as they relate to your results.
    b. Repeat part (a) for the four quarters of 2005, and again for the period from 2008:Q3 to 2009:Q2. Comment on the relationships among house prices, stock prices, and consumption as they relate to your results, before and during the crisis.
    c. How do the current household data compare to the data from the period prior to the financial crisis and during the crisis? Do you think the current data are indicative of a bubble?
2. **Real-Time Data Analysis** Go to the St. Louis Federal Reserve FRED database and find data on corporate net worth of nonfinancial businesses (TNWMVBSNNCB), private domestic investment (GPDIC1), and a measure of financial frictions, the St. Louis Fed financial stress index (STLFSI2). For all three measures, be sure to convert the *frequency* setting to "Quarterly." Download the data into a spreadsheet, and make sure the data align

correctly with the appropriate dates. For corporate net worth and private domestic investment, calculate the annualized growth rates from quarter to quarter. To do this, take the current-period data minus the previous-quarter data, then divide by the previous-quarter data. Multiply by 100 to change the results to percentage form, and then multiply by 4 to annualize the data.

a. Calculate the average growth rates over the most recent four quarters of data available for the corporate net worth and private domestic investment variables. Calculate the difference between the value of the stress index during the most recent quarter and the value of the stress index one year earlier. Comment on the relationships among financial stress, net wealth of corporate businesses, and private domestic investment.
b. Repeat part (a) for the four quarters of 2005 and for the period from 2008:Q3 to 2009:Q2. Comment on the relationships among financial stress, net wealth of corporate businesses, and private domestic investment before and during the crisis as they relate to your results. Assuming the financial stress measure is indicative of heightened asymmetric information problems, comment on how the crisis-period data relate to the typical dynamics of a financial crisis.
c. How do the current investment data compare to the data for the period prior to the financial crisis and during the crisis? Do you think the current data are indicative of a bubble?

PART 4

The Management of Financial Institutions

Crisis and Response: Canadian Banks Are the World's Best

During the global financial crisis, the value of mortgage-backed securities held by financial institutions plummeted, and governments in the United States and Europe worked on full-scale banking bailouts and rescue packages in the trillions of dollars. Canada, however, did not have to bail out any banks, although Canadian banks also had their problems. Their shares fell by almost 50%, and some of them experienced huge losses; for example, the CIBC suffered writedowns in excess of $6 billion in 2008 because of significant exposure to the U.S. capital and real estate markets.

One reason Canadian banks have done much better than their U.S. and European counterparts is because Canada's banking regulator, the Office of the Superintendent of Financial Institutions, has been more conservative than banking regulators in other countries. As a result, Canadian banks have lower leverage and more conservative lending and acquisition practices. In fact, in the aftermath of the Great Recession, Canada's banking system has been viewed as the soundest in the world, with many countries considering Canadian-style reforms of their financial markets.

Managing banks is not an easy task, and the global financial crisis highlights that it has become even more difficult because of increased complexity in the financial services marketplace. In Chapter 12 we look at the business and process of banking, and in Chapter 13 we study financial derivatives and how managers of financial institutions use derivatives to manage different types of risk.

12 Banking and the Management of Financial Institutions

Learning Objectives

12.1 Summarize the features of a bank balance sheet.

12.2 Apply changes to a bank's assets and liabilities on a T-account.

12.3 Identify the ways in which banks can manage their assets and liabilities to maximize profit.

12.4 List the ways in which banks deal with credit risk.

12.5 Apply gap analysis and duration analysis, and identify interest-rate risk.

12.6 Summarize the types of off-balance-sheet activities.

Preview

Because banking plays such a major role in channelling funds to borrowers with productive investment opportunities, this financial activity is important in ensuring that the financial system and the economy run smoothly and efficiently. In Canada and other countries, banks (depository institutions) provide loans to businesses, help us finance our college educations and our purchases of new cars and homes, and provide us with services such as chequing and savings accounts, debit cards, and ATMs.

In this chapter, we examine how banking is conducted to earn the highest profits possible: how and why banks make loans, how they acquire funds and manage their assets and liabilities (debts), and how they earn income. Although we focus on commercial banks, because they are the most important financial intermediary, many of the same principles are applicable to other types of financial intermediaries.

12.1 THE BANK BALANCE SHEET

LO 12.1 Summarize the features of a bank balance sheet.

To understand how banking works, we start by looking at the bank **balance sheet**, a list of the bank's assets and liabilities. As the name implies, this list balances; that is, it has the characteristic that

$$\text{total assets} = \text{liabilities} + \text{capital}$$

A bank's balance sheet is also a list of its *sources* of bank funds (liabilities and capital) and *uses* to which the funds are put (assets). Banks obtain funds by borrowing and by issuing other liabilities, such as deposits. They then use these funds to acquire assets such as securities and loans. Banks make profits by earning an interest rate on their holdings of securities and loans that is higher than the expenses on their liabilities. The balance sheet of all banks in Canada, as of December 31, 2020, appears in Table 12-1.

Liabilities

A bank acquires funds by issuing (selling) liabilities such as deposits, which are the *sources of funds* the bank uses. The funds obtained from issuing liabilities are used to purchase income-earning assets.

TABLE 12-1 Balance Sheet of All Banks in Canada (Items as a Percentage of the Total, December 31, 2020)

Assets (Uses of Funds)*		Liabilities (Sources of Funds)	
Reserves and cash items	10.83	Demand and notice deposits	39.74
Securities	21.86	Fixed-term deposits	26.65
Loans		Advances from the Bank of Canada	0.00
Non-mortgage loans	35.24	Borrowings	28.23
Mortgages	21.64	Bank capital	5.38
Other assets	10.43		
Total	100.00	Total	100.00

*In order of decreasing liquidity
Source: OSFI website, www.osfi-bsif.gc.ca.

Demand and Notice Deposits Demand deposits are payable on demand; that is, if a depositor shows up at the bank and requests payment by making a withdrawal, the bank must pay the depositor immediately. Similarly, if a person who receives a cheque written on an account from a bank presents that cheque at the bank, the bank must pay the funds out immediately (or credit them to that person's account). Notice deposits are more important as a source of funds for the banks than are demand deposits. Although notice deposits have a notice requirement in the contractual agreement with the client, the banks rarely enforce this clause, and so in fact most notice deposits are really just like demand deposits in this sense.

Demand deposits and notice deposits are bank accounts that allow the owner to write cheques to third parties. Table 12-1 shows that this category of chequable deposits is an important low-cost source of bank funds, making up close to 40% of bank liabilities. Once, chequable deposits were the most important source of bank funds, but with the appearance of new, more attractive financial instruments, the share of chequable deposits in total bank liabilities has shrunk over time.

A chequable deposit is an asset for the depositor because it is part of their wealth. Because the depositor can withdraw funds and the bank is obliged to pay, chequable deposits are a liability for the bank. They are usually the lowest-cost source of bank funds because depositors are willing to forgo some interest in order to have access to a liquid asset that can be used to make purchases. The bank's costs of maintaining chequable deposits include interest payments and the costs incurred in servicing these accounts—processing and storing cancelled cheques, preparing and sending out monthly statements, providing efficient tellers (human or otherwise), maintaining an impressive building and conveniently located branches, and advertising and marketing to entice customers to deposit their funds with a given bank.

Fixed-Term Deposits Fixed-term deposits are an important source of bank funds (about 27% of bank liabilities in Table 12-1). Owners cannot write cheques on fixed-term

deposits, but the interest rates paid on these deposits are usually higher than those on chequable deposits. There are two main types of fixed-term deposits: savings accounts and time deposits (also called *certificates of deposit*, or *CDs*).

Savings accounts were once the most common type of fixed-term deposit. Funds can be added to or withdrawn from savings accounts at any time. For these accounts, transactions and interest payments are recorded in a monthly statement or in a passbook held by the owners of the account.

Time deposits have a fixed maturity length, ranging from several months to over five years, and assess substantial penalties (the forfeiture of several months' interest) for early withdrawal of funds. Small-denomination time deposits (deposits of less than $100 000) are less liquid for the depositor than passbook savings, earn higher interest rates, and are a more costly source of funds for the banks.

Large-denomination time deposits (CDs) are available in denominations of $100 000 or over and are typically bought by corporations or other banks. Large-denomination CDs are negotiable; like bonds, they can be resold in a secondary market before they mature. For this reason, negotiable CDs are held by corporations, money market mutual funds, and other financial institutions as alternative assets to Treasury bills and other short-term bonds. Since 1964, when they first appeared in Canada, negotiable CDs have become an important source of bank funds.

Borrowings Banks obtain funds by borrowing from the Bank of Canada, other banks, and corporations. Borrowings from the Bank of Canada are called **overdraft loans** (also known as **advances**). Banks also borrow reserves overnight in the overnight market from other banks and financial institutions. Banks borrow funds overnight in order to have enough **settlement balances** at the Bank of Canada to facilitate the clearing of cheques and other transfers (these clearing and settlement processes will be discussed in detail in Chapter 16). Other sources of borrowed funds are loan arrangements with corporations (such as repurchase agreements) and borrowings of Eurodollars (deposits denominated in dollars residing in foreign banks or foreign branches of Canadian banks). Borrowings have become a more important source of bank funds over time: In 1960, they made up only a small fraction of bank liabilities; currently, they are about 28% of bank liabilities.

Bank Capital The final category on the liabilities side of the balance sheet is bank capital, the bank's net worth, which equals the difference between total assets and liabilities (about 5% of total bank assets in Table 12-1). Bank capital is raised by selling new equity (stock) or from retained earnings. A bank's capital is its cushion against insolvency, which occurs when a bank has liabilities in excess of assets, meaning that the bank can be forced into liquidation.

Assets

A bank uses the funds that it has acquired by issuing liabilities to purchase income-earning assets. Bank assets are thus naturally referred to as *uses of funds,* and the interest payments earned on them are what enable banks to make profits.

Reserves All banks hold some of the funds they acquire as deposits in an account at the Bank of Canada, in the form of settlement balances. **Reserves** consist of these settlement balances plus currency that is physically held by banks (called **vault cash** because it is stored in bank vaults overnight). Although Canadian banks are not

required to hold reserves in some proportion to their deposits (Canada removed all such legal requirements in June 1994) and there is a requirement of zero settlement balances with the Bank of Canada at the end of each banking day, banks hold some reserves, which we call **desired reserves**.

Banks hold reserves because of their desire to manage their own short-term liquidity requirements and respond to predictable clearing drains and predictable across-the-counter and automated banking machine drains. Moreover, banks hold reserves in order to meet unpredictable and potentially large withdrawals by their liability holders. The risk that net cash withdrawals might be negative is known as **banker's risk**, and from the perspective of this risk, banks hold reserves to meet unpredictable cash and clearing drains. We will refer to the fraction of deposits banks hold in the form of reserves as the **desired reserve ratio**.

Cash Items in Process of Collection Suppose that a cheque written on an account at another bank is deposited in your bank, and the funds for this cheque have not yet been received (collected) from the other bank. The cheque is classified as a "cash item in process of collection," and it is an asset for your bank because it is a claim on another bank for funds that will be paid within a few days.

Deposits at Other Banks Many small banks hold deposits in larger banks in exchange for a variety of services, including cheque collection, foreign exchange transactions, and help with securities purchases. This is one aspect of a system called *correspondent banking*.

Collectively, cash items in process of collection and deposits at other banks are referred to as *cash items*. In Table 12-1 they constitute about 11% of total assets, and their importance has been shrinking over time: In 1960, for example, they accounted for over 20% of total assets.

Securities A bank's holdings of securities are an important income-earning asset: Securities (made up entirely of debt instruments for commercial banks, because banks are not allowed to hold stock) account for about 22% of bank assets in Table 12-1. These securities can be classified into three categories: Government of Canada, provincial and municipal securities, and other securities. The Government of Canada securities are the most liquid because they can be easily traded and converted into cash with low transaction costs. Because of their high liquidity, short-term Canadian government securities are called **secondary reserves**.

Banks hold provincial and municipal government securities because provincial and municipal governments are more likely to do business with banks that hold their securities. Provincial and municipal government and other securities are both less marketable (less liquid) and riskier than Government of Canada securities, primarily because of default risk: Some possibility exists that the issuer of the securities may not be able to make its interest payments or pay back the face value of the securities when they mature.

Loans Banks make their profits primarily by issuing loans. In Table 12-1, some 57% of bank assets are in the form of loans, and in recent years they have generally produced more than half of bank revenues. A loan is a liability for the individual or corporation receiving it but an asset for a bank, because it provides income to the bank. Loans are typically less liquid than other assets because they cannot be turned into cash

until the loan matures. If a bank makes a one-year loan, for example, it cannot get its funds back until the loan comes due in one year. Loans also have a higher probability of default than other assets. Because of the lack of liquidity and higher default risk, the bank earns its highest return on loans.

As indicated in Table 12-1, the largest categories of loans for banks are non-mortgage loans (commercial and industrial loans made to businesses, consumer loans, and inter-bank loans) and mortgages. The balance sheet in Table 12-1 shows that non-mortgage loans exceed mortgage loans (about 35% versus 22% of bank assets). This is so because a larger portion of the foreign activities of Canadian banks is in corporate loans rather than mortgages. In fact, the major difference in the balance sheets of the various depository institutions is primarily in the type of loan they specialize in. Trust and loan companies and credit unions and *caisses populaires*, for example, specialize in residential mortgages.

Other Assets Bankers' acceptances and the physical capital (bank buildings, computers, and other equipment) owned by the banks are included in this category.

12.2 BASIC BANKING

LO 12.2 Apply changes to a bank's assets and liabilities on a T-account.

Before proceeding to a more detailed study of how a bank manages its assets and liabilities to make the highest profit, you should understand the basic operation of a bank.

In general terms, banks make profits by selling liabilities with one set of characteristics (a particular combination of liquidity, risk, size, and return) and using the proceeds to buy assets with a different set of characteristics. This process is often referred to as **asset transformation**. For example, a savings deposit held by one person can provide the funds that enable the bank to make a mortgage loan to another person. The bank has, in effect, transformed the savings deposit (an asset held by the depositor) into a mortgage loan (an asset held by the bank). Another way of describing this process of asset transformation is to say that the bank "borrows short and lends long" because it makes long-term loans and funds them by issuing short-term deposits.

The process of transforming assets and providing a set of services (cheque clearing, record keeping, credit analysis, and so forth) is like any other production process in a firm. If the bank produces desirable services at low cost and earns substantial income on its assets, it earns profits; if not, the bank suffers losses.

To make our analysis of the operation of a bank more concrete, we use a tool called a **T-account**. A T-account is a simplified balance sheet, with lines in the form of a T, that lists only the changes that occur in balance sheet items starting from some initial balance sheet position. Let's say that Jane Brown has heard that the First National Bank provides excellent service, so she opens a chequing account with a $100 bill. She now has a $100 chequable deposit at the bank, which shows up as a $100 liability on the bank's balance sheet. The bank now puts her $100 bill into its vault so that the bank's assets rise by the $100 increase in vault cash. The T-account for the bank looks like this:

First National Bank

Assets		Liabilities	
Vault cash	+$100	Chequable deposits	+$100

Because vault cash is also part of the bank's reserves, we can rewrite the T-account as follows:

Assets		Liabilities	
Reserves	+$100	Chequable deposits	+$100

Note that Jane Brown's opening of a chequing account leads to ***an increase in the bank's reserves equal to the increase in chequable deposits***.

If Jane had opened her account with a $100 cheque written on an account at another bank, say, the Second National Bank, we would get the same result. The initial effect on the T-account of the First National Bank would be as follows:

Assets		Liabilities	
Cash items in process of collection	+$100	Chequable deposits	+$100

Chequable deposits increase by $100 as before, but now the First National Bank is owed $100 by the Second National Bank. This asset for the First National Bank is entered in the T-account as $100 of cash items in process of collection because the First National Bank will now try to collect the funds that it is owed. It could go directly to the Second National Bank and ask for payment of the funds, but if the two banks are in separate states, that would be a time-consuming and costly process. Instead, the First National Bank deposits the cheque in its account at the Bank of Canada, and the Bank of Canada collects the funds from the Second National Bank. The net result is that the Bank of Canada transfers $100 of reserves from the Second National Bank to the First National Bank, and the final balance sheet positions of the two banks are as follows:

First National Bank

Assets		Liabilities	
Reserves	+$100	Chequable deposits	+$100

Second National Bank

Assets		Liabilities	
Reserves	–$100	Chequable deposits	–$100

The process initiated by Jane Brown can be summarized as follows: When a cheque written on an account at one bank is deposited in another, the bank receiving the deposit gains reserves equal to the amount of the cheque, while the bank on which the cheque is written sees its reserves fall by the same amount. Therefore, ***when a bank receives additional deposits, it gains an equal amount of reserves; when it loses deposits, it loses an equal amount of reserves.***

Now that you understand how banks gain and lose reserves, we can examine how a bank rearranges its balance sheet to make a profit when it experiences a change in its deposits. Let's return to the situation in which the First National Bank has just received

the extra $100 of chequable deposits. As you know, the bank wants to keep a certain fraction of its chequable deposits as desired reserves. If the fraction (the desired reserve ratio) is 10%, the First National Bank's desired reserves have increased by $10, and we can rewrite its T-account as follows:

Assets		Liabilities	
Desired reserves	+$10	Chequable deposits	+$100
Excess reserves	+$90		

Let's see how well the bank is doing as a result of the additional chequable deposits. Servicing the extra $100 of chequable deposits is costly, because the bank must keep records, pay tellers, pay for cheque clearing, and so forth. Because reserves earn little interest, the bank is taking a loss! The situation is even worse if the bank makes interest payments on the deposits. To make a profit, the bank must put to productive use all or part of the $90 of excess reserves it has available. One way to do this is to invest in securities. The other is to make loans; as we have seen, loans account for approximately 57% of the total value of bank assets (uses of funds). Because lenders are subject to the asymmetric information problems of adverse selection and moral hazard (discussed in Chapter 8), banks take steps to reduce the incidence and severity of these problems. Bank loan officers evaluate potential borrowers using what are called the "five C's"—character, capacity (ability to repay), collateral, conditions (in the local and national economies), and capital (net worth)—before they agree to lend. (A more detailed discussion of the methods banks use to reduce the risk involved in lending is given later in this chapter.)

Let's assume that the First National Bank chooses not to hold any excess reserves but to make loans instead. The T-account then looks like this:

First National Bank

Assets		Liabilities	
Desired reserves	+$10	Chequable deposits	+$100
Loans	+$90		

The bank is now making a profit because it holds short-term liabilities, such as chequable deposits, and uses the proceeds to fund longer-term assets, such as loans with higher interest rates. As mentioned earlier, this process of asset transformation is frequently described by saying that banks are in the business of "borrowing short and lending long." For example, if the loans have an interest rate of 10% per year, the bank earns $9 in income from its loans over the year. If the $100 of chequable deposits is in an account with a 5% interest rate and it costs another $3 per year to service the account, the cost per year of these deposits is $8. The bank's profit on the new deposits is then $1 per year, plus any interest that it earns on desired reserves.

12.3 GENERAL PRINCIPLES OF BANK MANAGEMENT

LO 12.3 Identify the ways in which banks can manage their assets and liabilities to maximize profit.

Now that you have some idea of how a bank operates, let's look at how a bank manages its assets and liabilities to earn the highest possible profit. The bank manager has four primary concerns. The first is to make sure that the bank has enough ready cash to pay its depositors when there are **deposit outflows**—that is, when deposits are lost because depositors make withdrawals and demand payment. To keep enough cash on hand, the bank must engage in **liquidity management**, the acquisition of assets that are liquid enough to meet the bank's obligations to depositors. Second, the bank manager must pursue an acceptably low level of risk by acquiring assets that have a low rate of default and by diversifying asset holdings (**asset management**). The manager's third concern is acquiring funds at low cost (**liability management**). Finally, the manager must decide the amount of capital the bank should maintain and then acquire the needed capital (**capital adequacy management**).

To understand bank (and other financial institution) management fully, we must go beyond the general principles of bank asset and liability management described next and look in more detail at how a financial institution manages its assets. The two sections following this one provide an in-depth discussion of how a financial institution manages **credit risk**, the risk arising because borrowers may default, and how it manages **interest-rate risk**, the riskiness of earnings and returns on bank assets caused by interest-rate changes.

Liquidity Management and the Role of Reserves

Let's see how a typical bank, the First National Bank, can deal with deposit outflows that occur when its depositors withdraw cash from chequing or savings accounts or write cheques that are deposited in other banks. In the example that follows, we assume that the bank has ample excess reserves and that all deposits have the same desired reserve ratio of 10% (the bank wants to keep 10% of deposits as reserves). Suppose that the First National Bank's initial balance sheet is as follows:

Assets		Liabilities	
Reserves	$20 million	Deposits	$100 million
Loans	$80 million	Bank capital	$ 10 million
Securities	$10 million		

The bank's desired reserves are 10% of $100 million, or $10 million. Given that it holds $20 million of reserves, the First National Bank has excess reserves of $10 million.

If a deposit outflow of $10 million occurs, the bank's balance sheet becomes

Assets		Liabilities	
Reserves	$10 million	Deposits	$90 million
Loans	$80 million	Bank capital	$10 million
Securities	$10 million		

The bank loses $10 million of deposits ***and*** $10 million of reserves, but because its desired reserves are now 10% of only $90 million (that is, $9 million), its reserves still exceed this amount by $1 million. In short, ***if a bank has ample excess reserves, a deposit outflow does not necessitate changes in other parts of its balance sheet.***

The situation is quite different when a bank holds insufficient excess reserves. Let's assume that instead of initially holding $10 million in excess reserves, the First National Bank makes additional loans of $10 million so that it holds no excess reserves. Its initial balance sheet would then be

Assets		Liabilities	
Reserves	$10 million	Deposits	$100 million
Loans	$90 million	Bank capital	$ 10 million
Securities	$10 million		

When it suffers the $10 million deposit outflow, its balance sheet becomes

Assets		Liabilities	
Reserves	$ 0	Deposits	$90 million
Loans	$90 million	Bank capital	$10 million
Securities	$10 million		

After $10 million has been withdrawn from deposits and hence reserves, the bank has a problem: Its desired reserves are 10% of $90 million, or $9 million, but it has no reserves! To eliminate this shortfall, the bank has four basic options. One option is to acquire the reserves needed to meet a deposit outflow by borrowing them from other banks in the overnight funds market or by borrowing from corporations.[1] If the First National Bank acquires the $9 million shortfall in reserves by borrowing it from other banks or corporations, its balance sheet becomes

Assets		Liabilities	
Reserves	$ 9 million	Deposits	$90 million
Loans	$90 million	Borrowings from other banks or corporations	$ 9 million
Securities	$10 million		
		Bank capital	$10 million

[1]One way in which the First National Bank can borrow from other banks and corporations is by selling negotiable certificates of deposit. This method for obtaining funds is discussed in the section on liability management.

The cost of this activity is the interest rate on these borrowings, such as the overnight interest rate.

A second alternative is for the bank to sell some of its securities to help cover the deposit outflow. For example, it might sell $9 million of its securities and deposit the proceeds with the Bank of Canada, resulting in the following balance sheet:

Assets		Liabilities	
Reserves	$ 9 million	Deposits	$90 million
Loans	$90 million	Bank capital	$10 million
Securities	$ 1 million		

The bank incurs some brokerage and other transaction costs when it sells these securities. The Canadian government securities that are classified as secondary reserves are very liquid, so the transaction costs of selling them are quite modest. However, the other securities the bank holds are less liquid, and the transaction costs can be appreciably higher.

The bank has a third option for meeting a deposit outflow: It can acquire reserves by borrowing from the Bank of Canada. In our example, the First National Bank could leave its security and loan holdings the same and borrow $9 million in loans from the Bank of Canada. Its balance sheet would then be

Assets		Liabilities	
Reserves	$ 9 million	Deposits	$90 million
Loans	$90 million	Borrowings from the Bank of Canada	$ 9 million
Securities	$10 million		
		Bank capital	$10 million

The cost associated with advances from the Bank of Canada is the interest rate that must be paid to the Bank of Canada (called the **discount rate**).

As a last option, a bank can acquire the $9 million of reserves to meet the deposit outflow by reducing its loans by this amount and depositing the $9 million it then receives with the Bank of Canada, thereby increasing its reserves by $9 million. This transaction changes the balance sheet as follows:

Assets		Liabilities	
Reserves	$ 9 million	Deposits	$90 million
Loans	$81 million	Bank capital	$10 million
Securities	$10 million		

The First National Bank is once again in good shape because its $9 million of reserves satisfies the (desired) reserve requirement.

However, this process of reducing its loans is the bank's costliest way of acquiring reserves when a deposit outflow exists. If the First National Bank has numerous

short-term loans renewed at fairly short intervals, it can reduce its total amount of loans outstanding fairly quickly by *calling in* loans—that is, by not renewing some loans when they come due. Unfortunately for the bank, this is likely to antagonize the customers whose loans are not being renewed because they have not done anything to deserve such treatment. Indeed, they are likely to take their business elsewhere in the future, a very costly consequence for the bank.

A second method of reducing its loans is for the bank to sell them off to other banks. Again, this is very costly because other banks do not know how risky these loans are and so may not be willing to buy the loans at their full value. (This is an example of the lemons adverse selection problem described in Chapter 8.)

The foregoing discussion explains why banks hold excess reserves even though loans or securities earn a higher return. When a deposit outflow occurs, excess reserves enable the bank to escape the costs of (1) borrowing from other banks or corporations, (2) selling securities, (3) borrowing from the Bank of Canada, or (4) calling in or selling off loans. ***Reserves are insurance against the costs associated with deposit outflows. The higher the costs associated with deposit outflows, the more excess reserves a bank will want to hold.***

Just as you and I would be willing to pay an insurance company to insure us against a casualty loss such as the theft of a car, a bank is willing to pay the cost of holding excess reserves (the opportunity cost—that is, the earnings forgone by not holding income-earning assets such as loans or securities) to insure against losses due to deposit outflows. Because excess reserves, like insurance, have a cost, banks also take other steps to protect themselves; for example, they might shift their holdings of assets to more liquid securities (secondary reserves).

Asset Management

Now that you understand a bank's need for liquidity, we can examine the basic strategy a bank pursues in managing its assets. To maximize its profits, a bank must simultaneously seek the highest returns possible on loans and securities, reduce risk, and make adequate provisions for liquidity by holding liquid assets. Banks try to accomplish these three goals in four basic ways.

First, banks try to find borrowers who will pay high interest rates and are unlikely to default on their loans. They seek out loan business by advertising their borrowing rates and by approaching corporations directly to solicit loans. It is up to the bank's loan officer to decide if potential borrowers are good credit risks who will make interest and principal payments on time (i.e., banks engage in screening to reduce the adverse selection problem). Typically, banks are conservative in their loan policies; the default rate is usually less than 1%. It is important, however, that banks not be so conservative that they miss out on attractive lending opportunities that earn high interest rates.

Second, banks try to purchase securities with high returns and low risk.

Third, in managing their assets, banks must attempt to lower risk by diversifying. They accomplish this by purchasing many different types of assets (short- and long-term, Government of Canada, and municipal bonds) and approving many types of loans to a variety of customers. Banks that have not sufficiently sought the benefits of diversification often come to regret it later. For example, banks that had overspecialized in making loans to energy companies, real estate developers, or farmers suffered huge losses in the 1980s with the slump in energy, property, and farm prices. Indeed, many of these banks (e.g., the Canadian Commercial Bank and Northland Bank) went broke because they had "put too many eggs in one basket."

Finally, the bank must manage the liquidity of its assets so that it can meet deposit outflows without bearing huge costs. This means that it will hold liquid securities even if they earn a somewhat lower return than other assets. The bank must decide, for example, how much it should hold in reserves to avoid the costs associated with a deposit outflow. In addition, it will want to hold Government of Canada securities as secondary reserves so that even if a deposit outflow forces some costs on the bank, these will not be terribly high. Again, it is not wise for a bank to be too conservative. If it avoids all costs associated with deposit outflows by holding only excess reserves, the bank suffers losses, because reserves earn low interest, while the bank's liabilities are costly to maintain. The bank must balance its desire for liquidity against the benefits of the increased earnings that can be obtained from less liquid assets, such as loans.

Liability Management

Before the 1960s, liability management was a staid affair: For the most part, banks took their liabilities as fixed and spent their time trying to achieve an optimal mix of assets. There were two main reasons for the emphasis on asset management. First, a large part of bank funds were obtained through chequable (demand) deposits that did not pay any interest. Thus banks could not actively compete with one another for these deposits by paying interest on them, and so their amount was effectively a given for an individual bank. Second, because the markets for making overnight loans between banks were not well developed, banks rarely borrowed from other banks to meet their reserve needs.

Starting in the 1960s, however, large banks (called **money centre banks**) in key financial centres, such as New York, Chicago, and San Francisco, began to explore ways in which the liabilities on their balance sheets could provide them with reserves and liquidity. This move led to an expansion of overnight loan markets, such as the federal funds market in the United States and the overnight funds market in Canada, and the development of new financial instruments, such as negotiable CDs (first developed in 1961), which enabled money centre banks to acquire funds quickly.[2]

This new flexibility in liability management meant that banks could take a different approach to bank management. They no longer needed to depend on chequable deposits as the primary source of bank funds and as a result no longer treated their sources of funds (liabilities) as given. Instead, they aggressively set target goals for their asset growth and tried to acquire funds (by issuing liabilities) as they were needed.

For example, today, when a money centre bank finds an attractive loan opportunity, it can acquire funds by selling a negotiable CD. Or, if it has a reserve shortfall, it can borrow funds from another bank in the overnight funds market without incurring high transaction costs. The overnight funds market can also be used to finance loans. Because of the increased importance of liability management, most banks now manage both sides of the balance sheet together in an *asset–liability management committee*.

The greater emphasis on liability management explains some of the important changes over the past three decades in the composition of banks' balance sheets. While negotiable CDs and bank borrowings have greatly increased in importance as a source of bank funds in recent years, chequable deposits have decreased in importance. Newfound flexibility in liability management and the search for higher profits have also stimulated banks to increase the proportion of their assets held in loans, which earn higher income.

[2]Because small banks are not as well known as money centre banks and so might be a higher credit risk, they find it harder to raise funds in the negotiable CD market. Hence they do not engage nearly as actively in liability management.

Capital Adequacy Management

Banks have to make decisions about the amount of capital they need to hold for three reasons. First, bank capital helps prevent *bank failure*, a situation in which the bank cannot satisfy its obligations to pay its depositors and other creditors and so goes out of business. Second, the amount of capital held affects returns for the owners (equity holders) of the bank. Third, a minimum amount of bank capital (bank capital requirements) is required by regulatory authorities.

How Bank Capital Helps Prevent Bank Failure

Let's consider two banks with identical balance sheets, except that High Capital Bank has a ratio of capital to assets of 10%, while Low Capital Bank has a ratio of 4%.

High Capital Bank

Assets		Liabilities	
Reserves	$10 million	Deposits	$90 million
Loans	$90 million	Bank capital	$10 million

Low Capital Bank

Assets		Liabilities	
Reserves	$10 million	Deposits	$96 million
Loans	$90 million	Bank capital	$ 4 million

Suppose both banks get caught up in the euphoria of the housing market only to find later that $5 million of their housing loans have become worthless. When these bad loans are written off (valued at zero), the total value of assets declines by $5 million. As a consequence, bank capital, which equals total assets minus liabilities, also declines by $5 million. The balance sheets of the two banks now look like this:

High Capital Bank

Assets		Liabilities	
Reserves	$10 million	Deposits	$90 million
Loans	$85 million	Bank capital	$ 5 million

Low Capital Bank

Assets		Liabilities	
Reserves	$10 million	Deposits	$96 million
Loans	$85 million	Bank capital	–$ 1 million

High Capital Bank takes the $5 million loss in stride because its initial cushion of $10 million in capital means that it still has a positive net worth (bank capital) of $5 million after the loss. Low Capital Bank, however, is in big trouble. The value of its assets has fallen below that of its liabilities, and its net worth is now –$1 million. Because the bank has a negative net worth, it is insolvent: It does not have sufficient assets to pay off all holders of its liabilities. When a bank becomes insolvent, government regulators close the bank, its assets are sold off, and its managers are fired. Because the owners of Low Capital Bank will find their investment wiped out, they clearly would have preferred the bank to have had a large enough cushion of bank capital to absorb the losses, as was the case for High Capital Bank. We therefore see an important rationale for a bank to maintain a sufficient level of capital: ***A bank maintains bank capital to lessen the chance that it will become insolvent.***

How the Amount of Bank Capital Affects Returns to Equity Holders

Because owners of a bank must know whether their bank is being managed well, they

need good measures of bank profitability. A basic measure of bank profitability is the **return on assets (ROA)**, the net profit after taxes per dollar of assets:

$$\text{ROA} = \frac{\text{net profit after taxes}}{\text{assets}}$$

The return on assets provides information on how efficiently a bank is being run because it indicates how much profit is generated, on average, by each dollar of assets.

However, what the bank's owners (equity holders) care about most is how much the bank is earning on their equity investment. This information is provided by the other basic measure of bank profitability, the **return on equity (ROE)**, which is defined as the net profit after taxes per dollar of equity (bank) capital:

$$\text{ROE} = \frac{\text{net profit after taxes}}{\text{equity capital}}$$

There is a direct relationship between the return on assets (which measures how efficiently the bank is run) and the return on equity (which measures how well the owners are doing on their investment). This relationship is determined by the **equity multiplier (EM)**, or the amount of assets per dollar of equity capital:

$$\text{EM} = \frac{\text{assets}}{\text{equity capital}}$$

To see this, we note that

$$\frac{\text{net profit after taxes}}{\text{equity capital}} = \frac{\text{net profit after taxes}}{\text{assets}} \times \frac{\text{assets}}{\text{equity capital}}$$

which, using our definitions, yields

$$\text{ROE} = \text{ROA} \times \text{EM} \tag{1}$$

The formula in Equation 1 tells us what happens to the return on equity when a bank holds a smaller amount of capital (equity) for a given amount of assets. As we have seen, High Capital Bank initially has \$100 million of assets and \$10 million of equity, which gives it an equity multiplier of 10 (= \$100 million/\$10 million). Low Capital Bank, by contrast, has only \$4 million of equity, so its equity multiplier is higher, equalling 25 (= \$100 million/\$4 million). Suppose these banks have been equally well run so that they both have the same return on assets, 1%. The return on equity for High Capital Bank equals 1% × 10 = 10%, whereas the return on equity for Low Capital Bank equals 1% × 25 = 25%. The equity holders in Low Capital Bank are clearly a lot happier than the equity holders in High Capital Bank because they are earning more than twice as high a return. We now see why the owners of a bank may not want it to hold too much capital. ***Given the return on assets, the lower the bank capital, the higher the return for the owners of the bank.***

Trade-Off Between Safety and Returns to Equity Holders We now see that bank capital has both benefits and costs. Bank capital benefits the owners of a bank in that it makes their investment safer by reducing the likelihood of bankruptcy. But bank capital is costly because the higher it is, the lower will be the return on equity for a given return on assets. In determining the optimal amount of bank capital, managers must compare the benefit of maintaining higher capital (increased safety) with the cost of higher capital (the lower return on equity for bank owners).

In more uncertain times, when the possibility of large losses on loans increases, bank managers might want to hold more capital to protect the equity holders. Conversely, if the managers have confidence that loan losses won't occur, they might want to reduce the amount of bank capital, have a high equity multiplier, and thereby increase the return on equity.

Bank Capital Requirements Banks also hold capital because they are required to do so by regulatory authorities. Because of the high costs of holding capital (for the reasons just described), bank managers often want to hold less bank capital relative to assets than is required by the regulatory authorities. In this case, the amount of bank capital is determined by the bank capital requirements. We discussed the details of bank capital requirements and their important role in bank regulation in Chapter 9.

APPLICATION Strategies for Managing Bank Capital

Suppose that, as the manager of the First National Bank, you have to make decisions about the appropriate amount of bank capital to hold in your bank. Looking at the balance sheet of the bank, which, like that of High Capital Bank, has a ratio of bank capital to assets of 10% ($10 million of capital and $100 million of assets), you are concerned that the large amount of bank capital is causing the return on equity to be too low. You conclude that the bank has a capital surplus and should increase the equity multiplier to raise the return on equity. What should you do next?

To lower the amount of capital relative to assets and raise the equity multiplier, you can do any of three things: (1) You can reduce the amount of bank capital by buying back some of the bank's stock; (2) you can reduce the bank's capital by paying out higher dividends to its stockholders, thereby reducing the bank's retained earnings; or (3) you can keep bank capital constant but increase the bank's assets by acquiring new funds—say, by issuing CDs—and then seeking out loan business or purchasing more securities with these new funds. Because you think it will enhance your position with the stockholders, you decide to pursue the second alternative and raise the dividend on the First National Bank stock.

Now suppose that the First National Bank is in a situation similar to that of Low Capital Bank and has a ratio of bank capital to assets of 4%. You now worry that the bank is short on capital relative to assets because it does not have a sufficient cushion to prevent bank failure. To raise the amount of capital relative to assets, you now have the following three choices: (1) You can raise capital for the bank by having it issue equity (common stock); (2) you can raise capital by reducing the bank's dividends to shareholders, thereby increasing retained earnings that it can put into its capital account; or (3) you can keep capital at the same level but reduce the bank's assets by making fewer loans or by selling off securities and then using the proceeds to reduce the bank's liabilities. Suppose that raising bank capital is not easy to do at the current time because bank stock is selling at low prices or because shareholders will protest if their dividends are cut. Then you might have to choose the third alternative and shrink the size of the bank.

Our discussion of the strategies for managing bank capital at the First National Bank leads to the following conclusion, which deserves particular emphasis: ***A shortfall of bank capital is likely to lead a bank to reduce its assets and therefore is likely to cause a contraction in lending.*** In past years, many banks experienced capital shortfalls and had to restrict asset and lending growth. The important consequences of this decision for the credit markets are illustrated by the application that follows. ◆

APPLICATION

How a Capital Crunch Caused a Credit Crunch During the Global Financial Crisis

The dramatic slowdown in the growth of credit in the wake of the financial crisis that began in 2007 triggered a "credit crunch" during which credit was hard to get. As a result, the performance of the economy in 2008 and 2009 was very poor. What caused the credit crunch?

Our analysis of how a bank manages its capital indicates that the 2008–2009 credit crunch was caused, at least in part, by the capital crunch, in which shortfalls of bank capital led to slower credit growth.

A major boom and bust in the U.S. housing market led to huge losses for banks from their holdings of securities backed by residential mortgages. These losses reduced bank capital, which led to capital shortfalls: Banks had to either raise new capital or restrict asset growth by cutting back on lending. Banks did raise some capital, but with the growing weakness of the economy, raising new capital was extremely difficult, so banks also chose to tighten their lending standards and reduce lending. Both of these reactions to capital shortfalls helped produce a weak economy in 2008 and 2009. ◆

12.4 MANAGING CREDIT RISK

LO 12.4 List the ways in which banks deal with credit risk.

As noted in our earlier discussion of the general principles of asset management, banks and other financial institutions must make successful loans that are paid back in full (and so subject the institution to little credit risk) if they are to earn high profits. The economic concepts of adverse selection and moral hazard (discussed in Chapters 2 and 8) provide a framework for understanding the principles that financial institutions must follow if they are to reduce credit risk and make successful loans.[3]

Adverse selection in loan markets occurs because bad credit risks (those most likely to default on their loans) are the ones who usually line up for loans; in other words, those who are most likely to produce an *adverse* outcome are also the most likely to be *selected*. Borrowers with very risky investment projects have much to gain if their projects are successful, so they are the most eager to obtain loans. Clearly, however, they are the least desirable borrowers because of the greater possibility that they will be unable to pay back their loans.

Moral hazard exists in loan markets because borrowers may have incentives to engage in activities that are undesirable from the lender's point of view. In such situations, it is more likely that the lender will be subjected to the *hazard* of default. Once borrowers have obtained a loan, they are more likely to invest in high-risk investment projects—projects that pay high returns to the borrower if successful. The high risk associated with these investments, however, makes it less likely that these borrowers will be able to pay back their loans.

To be profitable, financial institutions must overcome the adverse selection and moral hazard problems that make loan defaults more likely. Financial institutions attempt to solve these problems by using a number of principles for managing credit

[3]Other financial intermediaries, such as insurance companies, pension funds, and finance companies, also make private loans, and the credit risk management principles we outline here apply to them as well.

risk: screening and monitoring, establishment of long-term customer relationships, loan commitments, collateral and compensating balance requirements, and credit rationing.

Screening and Monitoring

Asymmetric information is present in loan markets because lenders have less information about the investment opportunities and activities of borrowers than borrowers do. This situation leads banks and other financial institutions to perform two information-producing activities: screening and monitoring. Indeed, Walter Wriston, a former head of Citicorp, one of the largest bank corporations in the United States, was often quoted as stating that the business of banking is the production of information.

Screening Adverse selection in loan markets requires lenders to screen out the bad credit risks from the good ones so that loans are profitable to them. To accomplish effective screening, lenders must collect reliable information from prospective borrowers. Effective screening and information collection together form an important principle of credit risk management.

When you apply for a consumer loan (such as a car loan or a mortgage to purchase a house), the first thing you are asked to do is fill out forms that elicit a great deal of information about your personal finances. You are asked about your salary, your bank accounts and other assets (such as cars, insurance policies, and furnishings), and your outstanding loans; your record of loan, credit card, and charge account repayments; and the number of years you've worked and the names of your employers. You also are asked personal questions such as your age, marital status, and number of children. The lender uses this information to evaluate how good a credit risk you are by calculating your credit score, a statistical measure derived from your answers that is used to predict whether you are likely to have trouble making your loan payments. Deciding how good a risk you are cannot be entirely scientific, so the lender must also use judgment. The loan officer, whose job it is to decide whether you should be given the loan, might call your employer or talk to some of the personal references you have supplied. The officer might even make a judgment based on your demeanour or your appearance. (This is why most people dress neatly and conservatively when they go to a bank to apply for a loan.)

The process of screening and collecting information is similar when a financial institution makes a business loan. It collects information about the company's profits and losses (income) along with information about its assets and liabilities. The lender also has to evaluate the likely future success of the business. So, in addition to obtaining information such as sales figures, a loan officer might ask questions about the company's future plans, the purpose of the loan, and competition within the industry. The officer may even visit the company to obtain a firsthand look at its operations. The bottom line is that, whether they are considering making personal loans or business loans, bankers and other financial institutions need to be nosy.

Specialization in Lending One puzzling feature of bank lending is that a bank often specializes in lending to local firms or to firms in particular industries, such as energy. In one sense, this behaviour seems surprising because it means that the bank is not diversifying its portfolio of loans and thus is exposing itself to more risk. But from another perspective, such specialization makes perfect sense. The adverse selection problem requires the bank to screen out bad credit risks. It is easier for the bank to collect information about local firms and determine their creditworthiness than to collect

comparable information on firms that are farther away. Similarly, by concentrating its lending on firms in specific industries, the bank becomes more knowledgeable about these industries and is therefore better able to predict which firms will be able to make timely payments on their debt.

Monitoring and Enforcement of Restrictive Covenants Once a loan has been made, the borrower has an incentive to engage in risky activities that make it less likely that the loan will be paid off. To reduce this moral hazard, financial institutions must write provisions (restrictive covenants) into loan contracts that restrict borrowers from engaging in risky activities. By monitoring borrowers' activities to see whether they are complying with the restrictive covenants and by enforcing the covenants if they are not, lenders can make sure that borrowers are not taking on risks at their expense. The need for banks and other financial institutions to engage in screening and monitoring explains why they spend so much money on auditing and information-collecting activities.

Long-Term Customer Relationships

An additional way for banks and other financial institutions to obtain information about their borrowers is through long-term customer relationships, another important principle of credit risk management.

If a prospective borrower has had a chequing account, a savings account, or a loan with a bank over a long period of time, a loan officer can look at past activity on the accounts and learn quite a bit about the borrower. The balances in the chequing and savings accounts tell the banker how liquid the potential borrower is and at what time of year the borrower has a strong need for cash. A review of the cheques the borrower has written reveals the borrower's suppliers. If the borrower has borrowed previously from the bank, the bank has a record of the loan payments. Thus long-term customer relationships reduce the costs of information collection and make it easier to screen out bad credit risks.

The need for monitoring by lenders adds to the importance of long-term customer relationships. If the borrower has borrowed from the bank before, the bank has already established procedures for monitoring that customer. Therefore, the costs of monitoring long-term customers are lower than the costs of monitoring new customers.

Long-term relationships benefit customers as well as banks. A firm that has had a previous relationship with a bank will find it easier to obtain a loan from the bank at a low interest rate because the bank has an easier time determining if the prospective borrower is a good credit risk and therefore incurs fewer costs in monitoring the borrower.

A long-term customer relationship has another advantage for the bank. No bank can think of every contingency when it writes a restrictive covenant into a loan contract; there will always be risky borrower activities that are not ruled out. However, what if a borrower wants to preserve a long-term relationship with a bank because it will be easier to get future loans from the bank at low interest rates? The borrower then has the incentive to avoid risky activities that would upset the bank, even if restrictions on these risky activities are not specified in the loan contract. Indeed, if a bank doesn't like what a borrower is doing even when the borrower isn't violating any restrictive covenants, it has some power to discourage the borrower from such activity: The bank can threaten not to give the borrower new loans in the future. Long-term customer relationships therefore enable banks to deal with even unanticipated moral hazard contingencies.

Loan Commitments

Banks also create long-term relationships and gather information by issuing **loan commitments** to commercial customers. A loan commitment is a bank's commitment (for a specified future period of time) to provide a firm with loans up to a given amount at an interest rate that is tied to some market interest rate. The majority of commercial and industrial loans are made under the loan commitment arrangement. The advantage for the firm is that it has a source of credit when it needs it. The advantage for the bank is that the loan commitment promotes a long-term relationship, which in turn facilitates information collection. In addition, provisions in the loan commitment agreement require the firm to continually supply the bank with information about the firm's income, asset and liability position, business activities, and so on. A loan commitment arrangement is a powerful method for reducing the bank's costs of screening and information collection.

Collateral and Compensating Balances

Collateral requirements for loans are important credit risk management tools. Collateral, which is property promised to the lender as compensation if the borrower defaults, lessens the consequences of adverse selection because it reduces the lender's losses in the case of a loan default. It also reduces moral hazard because the borrower has more to lose from a default. If a borrower defaults on a loan, the lender can sell the collateral and use the proceeds to make up for its losses on the loan. One particular form of collateral required when a bank makes commercial loans is called **compensating balances**: A firm receiving a loan must keep a required minimum amount of funds in a chequing account at the bank. For example, a business getting a $10 million loan may be required to keep compensating balances of at least $1 million in its chequing account at the bank. This $1 million in compensating balances can then be taken by the bank to make up some of the losses on the loan if the borrower defaults.

In addition to serving as collateral, compensating balances increase the likelihood that a loan will be paid off. They do this by helping the bank monitor the borrower and consequently reduce moral hazard. Specifically, by requiring the borrower to use a chequing account at the bank, the bank can observe the firm's cheque payment practices, which may yield a great deal of information about the borrower's financial condition. For example, a sustained drop in the borrower's chequing account balance may signal that the borrower is having financial trouble, or account activity may suggest that the borrower is engaging in risky activities; perhaps a change in suppliers means the borrower is pursuing a new line of business. Any significant change in the borrower's payment procedures is a signal to the bank that it should make inquiries. Compensating balances therefore make it easier for banks to monitor borrowers more effectively and are an important credit risk management tool.

Credit Rationing

Another way in which financial institutions deal with adverse selection and moral hazard is through **credit rationing**: refusing to make loans even though borrowers are willing to pay the stated interest rate or even a higher rate. Credit rationing takes two forms. The first occurs when a lender refuses to make a loan *of any amount* to a borrower, even if the borrower is willing to pay a higher interest rate. The second occurs when a lender is willing to make a loan but restricts the size of the loan to less than the borrower would like.

Initially, you might be puzzled by the first type of credit rationing. After all, even if the potential borrower is a credit risk, why doesn't the lender just extend the loan but at a higher interest rate? The answer is that adverse selection prevents this from being

a wise course of action. Individuals and firms with the riskiest investment projects are exactly those that are willing to pay the highest interest rates. If a borrower took on a high-risk investment and succeeded, the borrower would become extremely rich. But a lender wouldn't want to make such a loan precisely because the credit risk is high; the likely outcome is that the borrower will *not* succeed and the lender will not be paid back. Charging a higher interest rate just makes adverse selection worse for the lender; that is, it increases the likelihood that the lender is lending to a bad credit risk. The lender would therefore rather not make any loans at a higher interest rate; instead, it would engage in the first type of credit rationing and would turn down loans.

Financial institutions engage in the second type of credit rationing to guard against moral hazard: They grant loans to borrowers, but loans that are not as large as the borrowers want. Such credit rationing is necessary because the larger the loan, the greater the benefits from moral hazard. If a bank gives you a $1000 loan, for example, you are likely to take actions that enable you to pay it back because you don't want to hurt your credit rating for the future. However, if the bank lends you $10 million, you are more likely to fly down to Rio to celebrate. The larger your loan, the greater your incentives to engage in activities that make it less likely that you will repay the loan. Because more borrowers repay their loans if the loan amounts are small, financial institutions ration credit by providing borrowers with smaller loans than they seek.

12.5 MANAGING INTEREST-RATE RISK

LO 12.5 Apply gap analysis and duration analysis, and identify interest-rate risk.

With the increased volatility of interest rates that occurred in the 1980s, banks and other financial institutions became more concerned about their exposure to interest-rate risk, the riskiness of earnings and returns that is associated with changes in interest rates. To see what interest-rate risk is all about, let's again take a look at the First National Bank, which has the following balance sheet:

First National Bank

Assets		Liabilities	
Rate-sensitive assets	$20 million	Rate-sensitive liabilities	$50 million
Variable-rate and short-term loans		Variable-rate CDs	
Short-term securities		Money market deposit accounts	
Fixed-rate assets	$80 million	Fixed-rate liabilities	$50 million
Reserves		Chequable deposits	
Long-term loans		Savings deposits	
Long-term securities		Long-term CDs	
		Equity capital	

A total of $20 million of First National Bank's assets are rate-sensitive, with interest rates that change frequently (at least once a year), and $80 million of its assets are fixed-rate, with interest rates that remain unchanged for a long period (over a year). On the liabilities side, the First National Bank has $50 million of rate-sensitive liabilities and

$50 million of fixed-rate liabilities. Suppose that interest rates rise by five percentage points on average, from 10% to 15%. The income on the assets increases by $1 million (= 5% × $20 million of rate-sensitive assets), while the payments on the liabilities increase by $2.5 million (= 5% × $50 million of rate-sensitive liabilities). The First National Bank's profits now decline by $1.5 million (= $1 million − $2.5 million). Conversely, if interest rates fall by five percentage points, similar reasoning tells us that the First National Bank's profits increase by $1.5 million. This example illustrates the following point: ***If a bank has more rate-sensitive liabilities than assets, a rise in interest rates will reduce bank profits, and a decline in interest rates will raise bank profits.***

Gap and Duration Analysis

The sensitivity of bank profits to changes in interest rates can be measured more directly using **gap analysis**, in which the amount of rate-sensitive liabilities is subtracted from the amount of rate-sensitive assets. In our example, this calculation (called the "gap") is −$30 million (= $20 million − $50 million). By multiplying the gap times the change in the interest rate, we can immediately obtain the effect on bank profits. For example, when interest rates rise by five percentage points, the change in profits is 5% × −$30 million, which equals −$1.5 million, as we saw.

The analysis we just conducted is known as *basic gap analysis*, and it can be refined in two ways. Clearly, not all assets and liabilities in the fixed-rate category have the same maturity. One refinement, the *maturity bucket approach*, is to measure the gap for several maturity subintervals, called *maturity buckets*, so that effects of interest-rate changes over a multiyear period can be calculated. The second refinement, called *standardized gap analysis*, accounts for the differing degrees of rate sensitivity among rate-sensitive assets and liabilities.

An alternative method for measuring interest-rate risk, called **duration analysis**, examines the sensitivity of the market value of the bank's total assets and liabilities to changes in interest rates. Duration analysis is based on what is known as Macaulay's concept of *duration*, which measures the average lifetime of a security's stream of payments.[4] Duration is a useful concept because it provides a good approximation of the sensitivity of a security's market value to a change in its interest rate:

$$\text{percent change in market value of security} \approx -\text{percentage-point change in interest rate} \times \text{duration in years}$$

where ≈ denotes "approximately equals."

Duration analysis involves using the average (weighted) duration of a financial institution's assets and of its liabilities to see how its net worth responds to a change in interest rates. Going back to our example of the First National Bank, suppose the average duration of the bank's assets is three years (that is, the average lifetime of the stream of payments is three years), whereas the average duration of its liabilities is two years.

[4]Algebraically, Macaulay's duration, D, is defined as

$$D = \sum_{\tau=1}^{N} \tau \frac{CP_\tau}{(1+i)^\tau} \Big/ \sum_{\tau=1}^{N} \frac{CP_\tau}{(1+i)^\tau}$$

where τ = time until cash payment is made
CP_τ = cash payment (interest plus principal) at time τ
i = interest rate
N = time to maturity of the security

For a more detailed discussion of duration gap analysis using the concept of Macaulay's duration, see the online appendix to this chapter.

In addition, the First National Bank has $100 million of assets and, say, $90 million of liabilities, so its bank capital is 10% of assets. With a five-percentage-point increase in interest rates, the market value of the bank's assets falls by 15% (= −5% × 3 years), a decline of $15 million on the $100 million of assets. However, the market value of the liabilities falls by 10% (= −5% × 2 years), a decline of $9 million on the $90 million of liabilities. The net result is that the net worth (the market value of the assets minus the liabilities) has declined by $6 million, or 6% of the total original asset value. Similarly, a five-percentage-point decline in interest rates increases the net worth of the First National Bank by 6% of the total asset value.

As our example makes clear, both duration analysis and gap analysis indicate that the First National Bank will suffer if interest rates rise but will gain if they fall. Duration analysis and gap analysis are thus useful tools for a manager of a financial institution who is concerned about its degree of exposure to interest-rate risk.

APPLICATION Strategies for Managing Interest-Rate Risk

Suppose that as manager of the First National Bank, you have done a duration and gap analysis for the bank. Now you need to decide which alternative strategies you should pursue to manage the interest-rate risk.

If you firmly believe that interest rates will fall in the future, you may be willing to forgo action because you know that the bank has more rate-sensitive liabilities than rate-sensitive assets and so will benefit from the expected interest-rate decline. However, you also realize that the First National Bank is subject to substantial interest-rate risk because there is always a possibility that interest rates will rise rather than fall. What should you do to eliminate this interest-rate risk? One thing you could do is to shorten the duration of the bank's assets to increase their rate sensitivity. Alternatively, you could lengthen the duration of the liabilities. By this adjustment of the bank's assets or liabilities, the bank's income will be less affected by interest-rate swings.

One problem with eliminating the First National Bank's interest-rate risk by altering the balance sheet is that doing so might be very costly in the short run. The bank may be locked into assets and liabilities of particular durations because of where its expertise lies. Fortunately, recently developed financial instruments known as financial derivatives—financial forwards and futures, options, and swaps—can help the bank reduce its interest-rate risk exposure but do not require the bank to rearrange its balance sheet. ◆

12.6 OFF-BALANCE-SHEET ACTIVITIES

LO 12.6 Summarize the types of off-balance-sheet activities.

Although asset and liability management has traditionally been the primary concern of banks, in the more competitive environment of recent years banks have been aggressively seeking out profits by engaging in off-balance-sheet activities.[5] **Off-balance-sheet activities** involve trading financial instruments and generating income from fees and loan

[5]Managers of financial institutions also need to know how well their banks are doing at any point in time. A second appendix to this chapter, which can be found online, discusses how bank performance is measured.

sales, activities that affect bank profits but do not appear on bank balance sheets. Indeed, off-balance-sheet activities have been growing in importance for banks: The income from these activities as a percentage of assets has increased more than fivefold since 1980.

Loan Sales

One type of off-balance-sheet activity that has grown in importance in recent years involves income generated by loan sales. A **loan sale**, also called a *secondary loan participation*, involves a contract that sells all or part of the cash stream from a specific loan and thereby removes the loan so that it is no longer an asset on the bank's balance sheet. Banks earn profits by selling loans for amounts that are slightly greater than the amounts of the original loans. Because the high interest rate on these loans makes them attractive, institutions are willing to buy them, even though the higher price means that they earn a slightly lower interest rate than the original interest rate on the loan, usually on the order of 0.15 percentage points.

Generation of Fee Income

Another type of off-balance-sheet activity involves the generation of income from fees that banks receive for providing specialized services to their customers, such as making foreign exchange trades on a customer's behalf, servicing a mortgage-backed security by collecting interest and principal payments and then paying them out, guaranteeing debt securities such as banker's acceptances (by which the bank promises to make interest and principal payments if the party issuing the security cannot), and providing backup lines of credit. There are several types of backup lines of credit. We have already mentioned the most important, the loan commitment, under which, for a fee, the bank agrees to provide a loan at the customer's request, up to a given dollar amount, over a specified period of time. Credit lines with "overdraft privileges" are also now available to bank depositors—these bank customers can write cheques in excess of their deposit balances and, in effect, write themselves a loan. Other lines of credit for which banks get fees include standby letters of credit to back up issues of commercial paper and other securities, and credit lines (called *note issuance facilities* and *revolving underwriting facilities*) for underwriting Euronotes, which are medium-term Eurobonds.

Off-balance-sheet activities involving guarantees of securities and backup credit lines increase the risk a bank faces. Even though a guaranteed security does not appear on a bank's balance sheet, it still exposes the bank to default risk: If the issuer of the security defaults, the bank is left holding the bag and must pay off the security's owner. Backup credit lines also expose the bank to risk because the bank may be forced to provide loans when it does not have sufficient liquidity or when the borrower is a very poor credit risk.

Trading Activities and Risk Management Techniques

We have already mentioned that banks' attempts to manage interest-rate risk have led them to trading in financial futures, options for debt instruments, and interest-rate swaps. Banks engaged in international banking also conduct transactions in the foreign exchange market. All transactions in these markets are off-balance-sheet activities because they do not have a direct effect on the bank's balance sheet. Although bank trading in these markets is often directed toward reducing risk or facilitating other bank business, banks also try to outguess the markets and engage in speculation. This speculation can be a very risky business and has led to bank insolvencies, the most dramatic being the failure of Barings, a British bank, in 1995.

Global Barings, Daiwa, Sumitomo, Société Générale, and JP Morgan Chase: Rogue Traders and the Principal–Agent Problem

The demise of Barings, a venerable British bank more than a century old, is a sad morality tale of how the principal–agent problem operating through a rogue trader can take a financial institution that has a healthy balance sheet one month and turn it into an insolvent tragedy the next.

In July 1992, Nick Leeson, Barings's new head clerk at its Singapore branch, began to speculate on the Nikkei, the Japanese version of the Dow Jones stock index. By late 1992, Leeson had suffered losses of $3 million, which he hid from his superiors by stashing the losses in a secret account. He even fooled his superiors into thinking he was generating large profits, thanks to a failure of internal controls at his firm, which allowed him to execute trades on the Singapore exchange *and* oversee the bookkeeping of those trades. (As anyone who runs a cash business, such as a bar, knows, there is always a lower likelihood of fraud if more than one person handles the cash. Similarly, in trading operations, you never mix management of the back room with management of the front room; this principle was grossly violated by Barings management.)

Things didn't get better for Leeson, who by late 1994 had suffered losses exceeding $250 million. In January and February 1995, Leeson bet the bank. On January 17, 1995, the day of the earthquake in Kobe, Japan, he lost $75 million, and by the end of the week he had lost more than $150 million. When the stock market declined on February 23, leaving him with a further loss of $250 million, he called it quits and fled Singapore. Three days later, he turned himself in at the Frankfurt airport. By the end of his wild ride, Leeson's losses, $1.3 billion in all, ate up Barings's capital and caused the bank to fail. Leeson was subsequently convicted and sent to jail in Singapore for his activities. He was released in 1999 and apologized for his actions.

Our asymmetric information analysis of the principal–agent problem explains Leeson's behaviour and highlights the danger of a management lapse like Barings's. Letting Leeson control both his own trades and the back room increased asymmetric information, because it reduced the principal's (Barings's) knowledge about Leeson's trading activities. This lapse increased the moral hazard incentive for him to take risks at the bank's expense, as he was now less likely to be caught. Furthermore, once he had experienced large losses, he had even greater incentives to take on even higher risk because, if his bets worked out, he could reverse his losses and keep in good standing with the company; whereas if his bets soured, he had little to lose because he was out of a job anyway. Indeed, the bigger his losses, the more he had to gain by bigger bets, which explains the escalation of the amount of his trades as his losses mounted. If Barings's managers had understood the principal–agent problem, they would have been more vigilant at finding out what Leeson was up to, and the bank might still be here today.

Unfortunately, Nick Leeson is no longer a rarity in the rogue traders' billionaires club, made up of those who have lost more than $1 billion. Over 11 years, Toshihide Iguchi, an officer in the New York branch of Daiwa Bank, also had control of both the bond trading operation and the back room bookkeeping activities, and he racked up $1.1 billion in losses over the period. In July 1995, Iguchi disclosed his losses to his superiors, but the management of the bank did not disclose them to its regulators. The result was that Daiwa was slapped with a $340 million fine and the bank was thrown out of the country by U.S. bank regulators.

Yasuo Hamanaka is another member of the billionaires club. In July 1996, he topped Leeson's and Iguchi's record, losing $2.6 billion for his employer, the Sumitomo Corporation, one of Japan's top trading companies. In January 2008, Jerome Kerviel's loss for his bank, Société Générale, set the all-time record for a rogue trader: His unauthorized trades cost the French bank $7.2 billion. In 2012, even the highly successful JP Morgan Chase bank experienced a trading loss of over $2 billion, caused by rogue trader Bruno Iksill, who was colourfully nicknamed "the London Whale."

The moral of these stories is that management of firms engaged in trading activities must reduce the principal–agent problem by closely monitoring their traders' activities, or the rogues' gallery will continue to grow.

Trading activities, although often highly profitable, are also highly risky because they make it easy for financial institutions and their employees to make huge bets quickly. The principal–agent problem, discussed in Chapter 8, is an especially severe problem for management of trading activities. Because a trader (the agent) has the opportunity to place large bets, whether they trade in bond markets, foreign exchange markets, or financial derivatives, they have an incentive to take on excessive risks: If their trading strategy leads to large profits, they are likely to receive a high salary and bonuses, but if they take large losses, the financial institution (the principal) will have to cover them. As the Barings Bank failure of 1995 so forcefully demonstrated, a trader subject to the principal–agent problem can take an institution that is quite healthy and drive it into insolvency very rapidly (see the Global box on the preceding page).

To reduce the principal–agent problem, managers of financial institutions must set up internal controls to prevent debacles like the one at Barings. Such controls include the complete separation of the people in charge of trading activities from those in charge of the bookkeeping for trades. In addition, managers must set limits on the total amount of traders' transactions and on the institution's risk exposure. Managers must also scrutinize risk assessment procedures using the latest computer technology.

One such method involves the value-at-risk approach. In this approach, the institution develops a statistical model with which it can calculate the maximum loss that its portfolio is likely to sustain over a given time interval, dubbed the value at risk, or VaR. For example, a bank might estimate that the maximum loss it would be likely to sustain over one day with a probability of 1 in 100 is $1 million; the $1 million figure is the bank's calculated value at risk. Another approach is called "stress testing." In this approach, a manager uses a computer model to project what would happen if a doomsday scenario occurred; that is, they look at the losses the institution would sustain in the event of an unusual combination of bad events. With the value-at-risk approach and stress testing, a financial institution can assess its risk exposure and take steps to reduce it.

Bank regulators have become concerned about the increased risk that banks are facing from their off-balance-sheet activities and, as we discussed in Chapter 9, are encouraging banks to pay increased attention to risk management. In addition, the Bank for International Settlements is developing additional bank capital requirements based on value-at-risk calculations for a bank's trading activities.

SUMMARY

1. The balance sheet of a commercial bank can be thought of as a list of the sources and uses of bank funds. A bank's liabilities are its sources of funds, which include chequable deposits, savings and time deposits, advances from the Bank of Canada, borrowings from other banks and corporations, and bank capital. A bank's assets are its uses of funds, which include reserves, cash items in process of collection, deposits at other banks, securities, loans, and other assets (mostly physical capital).
2. Banks make profits through the process of asset transformation: They borrow short (accept short-term deposits) and lend long (make long-term loans). When a bank takes in additional deposits, it gains an equal amount of reserves; when it pays out deposits, it loses an equal amount of reserves.
3. Although more liquid assets tend to earn lower returns, banks still desire to hold them. Specifically, banks hold excess and secondary reserves because they provide insurance against the costs of a deposit outflow. Banks manage their assets to maximize profits by seeking the highest returns possible on loans and securities while at the same time trying to lower risk and make adequate provisions for liquidity. Although liability management was once a staid affair, large (money centre) banks now

actively seek out sources of funds by issuing liabilities such as negotiable CDs or by actively borrowing from other banks and corporations. Banks manage the amount of capital they hold with the goals of preventing bank failure and meeting bank capital requirements set by the regulatory authorities. However, banks do not want to hold too much capital because by so doing they will lower the returns to equity holders.

4. The concepts of adverse selection and moral hazard explain many credit risk management principles involving loan activities: screening and monitoring, establishment of long-term customer relationships and loan commitments, collateral and compensating balances, and credit rationing.
5. With the increased volatility of interest rates that occurred in the 1980s, financial institutions became more concerned about their exposure to interest-rate risk. Gap and duration analyses tell a financial institution if it has more rate-sensitive liabilities than assets (in which case a rise in interest rates will reduce profits, and a fall in interest rates will raise profits). Financial institutions manage their interest-rate risk by modifying their balance sheets but can also use strategies involving financial derivatives.
6. Off-balance-sheet activities consist of trading financial instruments and generating income from fees and loan sales, all of which affect bank profits but are not visible on bank balance sheets. Because these off-balance-sheet activities expose banks to increased risk, bank management must pay particular attention to risk assessment procedures and internal controls to restrict employees from taking on too much risk.

MyLab Economics Make the grade with MyLab Economics: Select end-of-chapter questions can be found on MyLab Economics. You can practise them as often as you want, and most feature step-by-step guided instructions to help you find the right answer.

QUESTIONS

1. Why might a bank be willing to borrow funds from other banks at a higher rate than the rate at which it can borrow from the Bank of Canada?
2. Rank the following bank assets from most to least liquid:
   a. Commercial loans
   b. Securities
   c. Reserves
   d. Physical capital
3. The bank you own has the following balance sheet:

Assets		Liabilities	
Reserves	$ 75 million	Deposits	$500 million
Loans	$525 million	Bank capital	$100 million

If the bank suffers a deposit outflow of $50 million with a desired reserve ratio on deposits of 10%, what actions should you take?

4. If a deposit outflow of $50 million occurs, which balance sheet would a bank rather have initially, the balance sheet in Question 3 or the following balance sheet? Why?

Assets		Liabilities	
Reserves	$100 million	Deposits	$500 million
Loans	$500 million	Bank capital	$100 million

5. If no decent lending opportunity arises in the economy, and the central bank pays an interest rate on reserves that is similar to other low-risk investments, do you think banks will be willing to hold large amounts of excess reserves?
6. If the bank you own has no excess reserves and a sound customer comes in asking for a loan, should you automatically turn the customer down, explaining that you don't have any excess reserves to lend out? Why or why not? What options are available that will enable you to provide the funds your customer needs?

7. If a bank finds that its ROE is too low because it has too much bank capital, what can it do to raise its ROE?
8. If a bank is falling short of meeting its capital requirements by $1 million, what three things can it do to rectify the situation?
9. Why do equity holders care more about ROE than about ROA?
10. If a bank doubles the amount of its capital and ROA stays constant, what will happen to ROE?
11. What are the benefits and costs for a bank when it decides to increase the amount of its bank capital?
12. Why is being nosy a desirable trait for a banker?
13. A bank almost always insists that the firms it lends to keep compensating balances at the bank. Why?
14. If the president of a bank told you that the bank was so well run that it has never had to call in loans, sell securities, or borrow as a result of a deposit outflow, would you be willing to buy stock in that bank? Why or why not?
15. "Because diversification is a desirable strategy for avoiding risk, it never makes sense for a bank to specialize in making specific types of loans." Is this statement true, false, or uncertain? Explain your answer.
16. If you are a banker and expect interest rates to rise in the future, would you prefer to make short-term loans or long-term loans?
17. "Bank managers should always seek the highest return possible on their assets." Is this statement true, false, or uncertain? Explain your answer.
18. After July 2010, bank customers in the United States using a debit card had to specifically opt-in to the bank's overdraft protection plan. Explain the effect of this regulation on a bank's noninterest income.

APPLIED PROBLEMS

19. Using the T-accounts of the First National Bank and the Second National Bank given in this chapter, describe what happens when Jane Brown writes a $50 cheque on her account at the First National Bank to pay her friend Joe Green, who in turn deposits the cheque in his account at the Second National Bank.
20. What happens to reserves at the First National Bank if one person withdraws $1000 of cash and another person deposits $500 of cash? Use T-accounts to explain your answer.
21. Angus Bank holds no excess reserves. The desired reserves ratio is 9%, and reserves are currently $27 million. Determine the amount of deposits, the reserve shortage created by a deposit outflow of $5 million, and the cost of the reserve shortage if Angus Bank borrows in the overnight funds market (assume the overnight interest rate is 0.25%).
22. Excess reserves act as insurance against deposit outflows. Suppose that on a yearly basis Malcom Bank holds $12 million in excess reserves and $88 million in desired reserves. Suppose that Malcom Bank can earn 3.5% on its loans and that the interest paid on reserves is 0.2%. What would be the cost of this insurance policy?
23. Victory Bank reports an EM of 25, while Batovi Bank reports an EM equal to 14. Which bank is better prepared to respond against large losses on loans?
24. Suppose you are the manager of a bank whose $100 billion of assets have an average duration of four years and whose $90 billion of liabilities have an average duration of six years. Conduct a duration analysis for the bank, and show what will happen to the net worth of the bank if interest rates rise by two percentage points. What actions could you take to reduce the bank's interest-rate risk?
25. Suppose you are the manager of a bank that has $15 million of fixed-rate assets, $30 million of rate-sensitive assets, $25 million of fixed-rate liabilities, and $20 million of rate-sensitive liabilities. Conduct a gap analysis for the bank, and show what will happen to bank profits if interest rates rise by five percentage points. What actions could you take to reduce the bank's interest-rate risk?

DATA ANALYSIS PROBLEM

1. Go to the Office of the Superintendent of Financial Institutions (OSFI) and find data (as of December 31, 2020) for all domestic banks on total liabilities, total deposits, and residual of assets less liabilities.
    a. What is the balance sheet interpretation of the residual of assets less liabilities?
    b. Use the three indicators listed above to calculate the total amount of borrowings by banks.
2. Go to the OSFI website and find data for all domestic banks on total assets, reserves, securities held, commercial and industrial loans, and other assets. Use the data for December 31, 2020, across all indicators.
    a. What is the total amount of loans held by banks? What is this number as a percentage of total bank assets?
    b. What is the total amount of securities held by banks? What is this number as a percentage of total bank assets?
    c. What is the total amount of cash items held? What is this number as a percentage of total bank assets?

13 Risk Management with Derivatives

Learning Objectives

13.1 Define a hedge, a long position, and a short position.

13.2 Define a forward contract, and summarize its advantages and disadvantages.

13.3 Summarize the differences between a forward contract and a financial futures contract.

13.4 Identify the different types of options contracts, and summarize the three conclusions regarding call and put options.

13.5 Define a swap, and summarize the advantages and disadvantages of interest-rate swaps.

13.6 Summarize the three types of credit derivatives.

Preview

Starting in the 1970s, and increasingly in the 1980s and 1990s, the world became a riskier place for financial institutions. Swings in interest rates widened, and the bond and stock markets went through numerous episodes of increased volatility. As a result of these developments, managers of financial institutions became more concerned with reducing the risk their institutions faced. Given the greater demand for risk reduction, the process of financial innovation described in Chapter 10 came to the rescue by producing new financial instruments to help financial institution managers better manage risk. These instruments, called **financial derivatives**, have payoffs that are linked to previously issued securities and are extremely useful risk reduction tools.

In this chapter, we look at the most important financial derivatives that managers of financial institutions use to reduce risk: forward contracts, financial futures, options, and swaps. We examine not only how the markets for each of these financial derivatives work but also how they can be used by financial institutions to manage risk. We also study financial derivatives because they have become an important source of profits for financial institutions, particularly larger banks, which, as we saw in Chapter 10, have found their traditional business declining.

13.1 HEDGING

LO 13.1 Define a hedge, a long position, and a short position.

Financial derivatives are effective at reducing risk because they enable financial institutions to **hedge**—that is, to engage in a financial transaction that reduces or eliminates risk. When a financial institution has bought an asset, it is said to have taken a **long position**, and this exposes the institution to risk if the returns on the asset are uncertain. Conversely, if it has sold an asset that it has agreed to deliver to another party at a future date, it is said to have taken a **short position**, and this can also expose the institution to risk. Financial derivatives can be used to reduce risk by invoking the following basic principle of hedging: ***Hedging risk involves engaging in a financial transaction that offsets a long position by taking an additional short position, or offsets a short position by taking an additional long position***. In other words, if a financial institution has *bought* a security and has therefore taken a long position, it conducts a hedge by contracting to *sell* that security (take a short position) at some future date. Alternatively, if it has taken a short position by *selling* a security that it needs to deliver

at a future date, then it conducts a hedge by contracting to *buy* that security (take a long position) at a future date. We look at how this principle can be applied using forward and futures contracts.

13.2 INTEREST-RATE FORWARD CONTRACTS

LO 13.2 Define a forward contract, and summarize its advantages and disadvantages.

A **forward contract** is an agreement between two parties to engage in a financial transaction at a future (forward) point in time. Here we focus on forward contracts that are linked to debt instruments, called **interest-rate forward contracts**; later in the chapter, we discuss forward contracts for foreign currencies.

Interest-rate forward contracts involve the future sale (or purchase) of a debt instrument and have several dimensions: (1) specification of the actual debt instrument that will be delivered at a future date, (2) the amount of the debt instrument to be delivered, (3) the price (or interest rate) to be paid on the debt instrument when it is delivered, and (4) the date on which delivery will take place. An example of an interest-rate forward contract might be an agreement that the First National Bank will sell to the Rock Solid Insurance Company, one year from today, $5 million face value of the 6s of 2041 Canada bonds (that is, coupon bonds with a 6% coupon rate that mature in 2041) at a price that yields the same interest rate on these bonds as today's rate—say, 6%. Because Rock Solid will buy the securities at a future date, it is said to have taken a long position, whereas the First National Bank, which will sell the securities on that date, has taken a short position.

APPLICATION Hedging with Interest-Rate Forward Contracts

Why would the First National Bank want to enter into this forward contract with Rock Solid Insurance Company in the first place?

To understand, suppose that you are the manager of the First National Bank and that you have bought $5 million of the 6s of 2041 Canada bonds. The bonds are currently selling at par value, so their yield to maturity is 6%. Because these are long-term bonds, you recognize that you are exposed to substantial interest-rate risk: If interest rates rise in the future, the price of these bonds will fall, resulting in a substantial capital loss that may cost you your job. How do you hedge this risk?

Knowing the basic principle of hedging, you see that your long position in these bonds can be offset by an equal short position for the same bonds with a forward contract. That is, you need to contract to sell these bonds at a future date at the current par value price. To ensure that this happens, you form an agreement with another party—in this case, Rock Solid Insurance Company—in which you promise to sell to Rock Solid the $5 million of the 6s of 2041 Canada bonds at par one year from today. By entering into this forward contract, you have successfully hedged against interest-rate risk. By locking in the future price of the bonds, you have eliminated the price risk that you face from uncertain interest-rate changes.

Why would Rock Solid Insurance Company want to enter into the forward contract with the First National Bank? Rock Solid expects to receive premiums of $5 million in one year's time that it will want to invest in the 6s of 2041 but worries that interest rates on these bonds will decline between now and next year. By using the

forward contract, it is able to lock in the 6% interest rate on the Canada bonds that it will buy from the First National Bank.

Pros and Cons of Forward Contracts

The advantage of forward contracts is that they can be as flexible as the parties involved want them to be. This means that an institution like the First National Bank may be able to hedge completely the interest-rate risk on the exact security it is holding in its portfolio, just as it did in our example.

However, forward contracts suffer from two problems that severely limit their usefulness. The first is that it may be very hard for an institution like the First National Bank to find another party (called a *counterparty*) to make the contract with. There are brokers to facilitate the matching up of parties like the First National Bank and the Rock Solid Insurance Company, but few institutions may want to engage in a forward contract specifically for the 6s of 2041. This means that it may prove impossible to find a counterparty when a financial institution (like the First National Bank) wants to make a specific type of forward contract. Furthermore, even if the First National Bank finds a counterparty, it may not get the price it was hoping for because, if the counterparty is not willing to pay the asking price, there may not be anyone else available to make the deal with. A serious problem for the market in interest-rate forward contracts, then, is that it can be difficult for one party to find another party with which to make a particular financial agreement; even if a willing party is found, the transaction may have to be made at a disadvantageous price. In the parlance of financial economists, this market suffers from a *lack of liquidity*. (Note that this use of the term *liquidity,* in which it is applied to a market, is somewhat broader than its use when it is applied to an asset. For an asset, liquidity refers to the ease with which the asset can be turned into cash; for a market, liquidity refers to the ease of carrying out financial transactions.)

The second problem with forward contracts is that they are subject to default risk. Suppose that in one year's time, interest rates rise, and so the price of the 6s of 2041 falls. The Rock Solid Insurance Company might then decide that it would like to default on the forward contract with the First National Bank, because it can now buy the bonds at the lower price. Or perhaps Rock Solid may not have been rock solid after all and may have gone bust during the year, and therefore it is no longer available to complete the terms of the forward contract. Because no outside organization is guaranteeing the contract, the only recourse is for the First National Bank to go to the courts to sue Rock Solid, but this process will be costly. Furthermore, if Rock Solid is already bankrupt, the First National Bank will suffer a loss; the bank can no longer sell the 6s of 2041 at the price it had agreed on with Rock Solid but instead will have to sell at a price well below that, because the price of these bonds has fallen.

The presence of default risk in forward contracts means that parties to these contracts must check that the counterparty is both financially sound and likely to live up to its contractual obligations. Because this type of investigation is costly, and because all the adverse selection and moral hazard problems discussed in earlier chapters apply, default risk is a major barrier to the use of interest-rate forward contracts. If we consider the dual problems of default risk and lack of liquidity, we see that forward contracts may be of limited usefulness to financial institutions. Although a market does exist for interest-rate forward contracts, particularly in Canada and mortgage-backed securities, it is not nearly as large as the financial futures market, to which we turn next. ◆

13.3 FINANCIAL FUTURES CONTRACTS AND MARKETS

LO 13.3 Summarize the differences between a forward contract and a financial futures contract.

Given the default risk and liquidity problems inherent in the interest-rate forward market, another solution to hedging interest-rate risk was needed. The solution was provided by the Chicago Board of Trade when it began developing financial futures contracts in 1975.

A **financial futures contract** is similar to an interest-rate forward contract in that it specifies that a financial instrument must be delivered by one party to another on a stated future date. However, it differs from an interest-rate forward contract in several ways, and these differences help overcome some of the liquidity and default problems of forward markets. In Canada, financial futures contracts are traded on the Montréal Exchange, which maintains active markets in short- and long-term Canada bond futures and stock index futures.

To understand what financial futures contracts are all about, let's look at one of the most widely traded futures contracts in the world—that for U.S. Treasury bonds, which are traded on the Chicago Board of Trade. The contract value is for $100 000 face value of bonds. Prices are quoted in points, with each point equal to $1000, and the smallest change in price is 1/32 of a point ($31.25). This contract specifies that the bonds to be delivered must have at least 15 years to maturity at the delivery date (and also must not be *callable*—that is, redeemable by the U.S. Treasury at its option—in less than 15 years). If the Treasury bonds delivered to settle the futures contract have a coupon rate different from the (for example) 6% specified in the futures contract, the amount of bonds to be delivered is adjusted to reflect the difference in value between the delivered bonds and the 6% coupon bond. In line with the terminology used for forward contracts, parties who have bought a futures contract and thereby have agreed to buy (take delivery of) the bonds are said to have taken a *long position*, and parties who have sold a futures contract and thereby have agreed to sell (deliver) the bonds have taken a *short position*.

To make our understanding of this type of contract more concrete, let's consider what happens when you buy or sell a U.S. Treasury bond futures contract. Let's say that on February 1, you sell one $100 000 June contract at a price of 115 (that is, $115 000). By selling this contract, you agree to deliver $100 000 face value of the long-term Treasury bonds to the contract's counterparty at the end of June for $115 000. By buying the contract at a price of 115, the buyer has agreed to pay $115 000 for the $100 000 face value of bonds when you deliver them at the end of June. If interest rates on long-term bonds rise, so that when the contract matures at the end of June the price of these bonds has fallen to 110 ($110 000 per $100 000 of face value), the buyer of the contract will have lost $5000, because they paid $115 000 for the bonds but can sell them only for the market price of $110 000. But you, the seller of the contract, will have gained $5000, because you can now sell the bonds to the buyer for $115 000 even though you have to pay only $110 000 for them in the market.

It is even easier to describe what happens to the parties who have purchased futures contracts and those who have sold futures contracts if we recognize the following fact: ***At the expiration date of a futures contract, the price of the contract converges to the price of the underlying asset to be delivered.*** To see why this is the case, consider what happens on the expiration date of the contract at the end of June when the price of the underlying $100 000 face value Treasury bond is 110 ($110 000). If the futures contract is selling below 110—say, at 109—a trader can buy the contract for $109 000, take delivery of the bond, and immediately sell it for $110 000, thereby earning a quick

profit of $1000. Because earning this profit involves no risk, it is a great deal that everyone would like to get in on. That means that everyone will try to buy the contract, and as a result, its price will rise. Only when the price rises to 110 will the profit opportunity cease to exist and the buying pressure disappear. Conversely, if the price of the futures contract is above 110—say, at 111—everyone will want to sell the contract. Now the sellers get $111 000 from selling the futures contract but have to pay only $110 000 for the Treasury bonds that they must deliver to the buyer of the contract, and the $1000 difference is their profit. Because this profit involves no risk, traders will continue to sell the futures contract until its price falls back down to 110, at which price there are no longer any profits to be made. The elimination of riskless profit opportunities in the futures market is referred to as **arbitrage**, and it guarantees that the price of a futures contract at expiration equals the price of the underlying asset to be delivered.[1]

Armed with our knowledge that a futures contract at expiration equals the price of the underlying asset, we can see even more easily who profits and who loses from such a contract when interest rates change. If interest rates have risen, so that the price of the U.S. Treasury bond is 110 on the expiration day at the end of June, the June Treasury bond futures contract will also have a price of 110. Thus, if you bought the contract for 115 in February, you have suffered a loss of 5 points, or $5000 (5% of $100 000). But if you sold the futures contract at 115 in February, the decline in price to 110 means that you have gained a profit of 5 points, or $5000.

APPLICATION Hedging with Financial Futures

The First National Bank can also use financial futures contracts to hedge the interest-rate risk on its holdings of $5 million of the 6s of 2041. To see how, suppose that in March 2019, the 6s of 2041 are the long-term bonds that will be delivered in the Montréal Exchange's Canada bond futures contract expiring one year in the future, in March 2020.[2] Also suppose that the interest rate on these bonds is expected to remain at 6% over the next year so that both the 6s of 2041 and the futures contract are selling at par (i.e., the $5 million of bonds is selling for $5 million and the $100 000 futures contract is selling for $100 000). The basic principle of hedging indicates that you, as the manager of the First National Bank, need to offset the long position in these bonds with a short position, so you have to sell the futures contract. But how many contracts should you sell? The number of contracts required to hedge the interest-rate risk is found by dividing the dollar value of the asset to be hedged by the dollar value of each contract, as shown in Equation 1:

$$NC = VA/VC \tag{1}$$

where NC = number of contracts for the hedge
VA = value of the asset
VC = value of each contract.

Given that the 6s of 2041 are the long-term bonds that will be delivered in the Montréal Exchange's Canada bond futures contract expiring one year in the future and that the

[1]In actuality, futures contracts sometimes set conditions for the timing and delivery of the underlying assets that cause the price of the contract at expiration to differ slightly from the price of the underlying assets. Because the difference in price is extremely small, we ignore it in this chapter.

[2]In the real world, designing a hedge is somewhat more complicated than indicated by the example here, because the bond that is most likely to be delivered might not be a 6s of 2041.

interest rate on these bonds is expected to remain at 6% over the next year so that both the 6s of 2041 and the futures contract are selling at par, how many contracts must First National sell to remove its interest-rate risk exposure from its $5 million holdings of the 6s of 2041? Since VA = $5 million and VC = $100 000,

$$\text{NC} = \$5 \text{ million}/\$100\,000 = 50.$$

You therefore hedge the interest-rate risk by selling 50 of the Canada bond futures contracts.

Now suppose that over the next year, an increased threat of inflation causes interest rates to increase to 8%. The value of the 6s of 2041 that the First National Bank is holding will then fall to $4 039 640 in March 2020.[3] Thus the loss from the long position in these bonds is $960 360:

Value on March 2020 @ 8% interest rate	$4 039 640
Value on March 2019 @ 6% interest rate	−$5 000 000
Loss	−$ 960 360

However, the short position in the 50 futures contracts that obligate you to deliver $5 million of the 6s of 2041 on March 2020 has a value equal to $4 039 640, the value of the $5 million of bonds after the interest rate has risen to 8%, as we saw before. Yet when you sold the futures contract, the buyer was obligated to pay you $5 million on the maturity date. Thus the gain from the short position on these contracts is also $960 360:

Amount paid to you on March 2020, agreed upon in March 2019	$5 000 000
Value of bonds delivered on March 2020 @ 8% interest rate	−$4 039 640
Gain	$ 960 360

Therefore, the net gain for the First National Bank is zero, indicating that the hedge has been conducted successfully. ◆

The hedge just described is called a **micro hedge** because the financial institution is hedging the interest-rate risk for a specific asset it is holding. A second type of hedge that financial institutions engage in is called a **macro hedge**, in which the hedge is for the institution's entire portfolio. For example, if a bank has more rate-sensitive liabilities than assets, we learned in Chapter 12 that a rise in interest rates will cause the value of the bank's net worth to decline. By selling interest-rate futures contracts that will yield a profit when interest rates rise, the bank can offset the losses on its overall portfolio from an interest-rate rise and thereby hedge its interest-rate risk.

[3]The value of the bonds can be calculated using a financial calculator as follows: FV = $5 000 000, PMT = $300 000, I = 8%, N = 19, PV = $4 039 640.

Organization of Trading in Financial Futures Markets

Financial futures contracts are traded on organized exchanges such as the Chicago Board of Trade, the Chicago Mercantile Exchange, the Montréal Exchange, and the London International Financial Futures Exchange. These futures exchanges are highly competitive with one another, and each organization tries to design contracts and set rules that will increase the amount of futures trading on its exchange. The exchanges are also regulated to ensure that prices in the market are not being manipulated. Table 13-1 lists some of the most popular financial futures contracts in the United States, along with their contract size and the number of contracts outstanding, called **open interest**, for July 23, 2020.

Given the globalization of other financial markets in recent years, it is not surprising that increased competition from abroad has been occurring in financial futures markets as well.

The Globalization of Financial Futures Markets

Because the United States was the first country to develop financial futures, American futures exchanges dominated the trading of financial futures in the early 1980s. For example, in 1985, all of the top 10 futures contracts were traded on exchanges in the

TABLE 13-1 Widely Traded Financial Futures Contracts in the United States

Type of Contract	Contract Size	Open Interest, July 23, 2020
Interest-Rate Contracts		
Treasury bonds	$100 000	1 066 642
Treasury notes	$100 000	3 474 435
Five-year Treasury notes	$100 000	3 466 672
Two-year Treasury notes	$200 000	2 110 429
Eurodollar	$4 million	9 680 371
Stock Index Contracts		
Standard & Poor's 500 Index	$250 × index	17 833
E-Mini Standard & Poor's 500 Index	$50 × index	2 641 558
E-Mini NASDAQ 100	$20 × index	223 911
Nikkei 225 Stock Average	$5 × index	37 143
Currency Contracts		
Yen	12 500 000 yen	142 745
Euro	125 000 euros	658 140
Canadian dollar	100 000 Canadian $	142 331
British pound	100 000 pounds	172 806
Swiss franc	125 000 francs	52 964
Mexican peso	500 000 new pesos	130 149

Source: Volume and Open Interest, www.cmegroup.com/market-data/volume-open-interest/. Used by permission of CME Group Inc.

United States. With the rapid growth of financial futures markets and the high profits made by the American futures exchanges, foreign exchanges saw a profit opportunity and began to enter this business. By the 1990s, Eurodollar contracts traded on the London International Financial Futures Exchange, Japanese government bond contracts and Euroyen contracts traded on the Tokyo Stock Exchange, French government bond contracts traded on the Marché à Terme International de France (now Euronext), and Nikkei 225 contracts traded on the Osaka Securities Exchange all were among the most widely traded futures contracts in the world.

International competition has spurred knockoffs of the most popular financial futures contracts initially developed in the United States. These contracts, which are traded on financial futures exchanges in other countries, are virtually identical to those traded in the United States and have the advantage that they can be traded when the American exchanges are closed. The movement to 24-hour-a-day trading in financial futures has been further stimulated by the development of the Globex electronic trading platform, which allows traders throughout the world to trade futures even when the exchanges are not officially open. Financial futures trading thus has become completely internationalized, and competition between U.S. and foreign exchanges is now intense.

Explaining the Success of Futures Markets

Futures contracts were designed to overcome the liquidity problem inherent in forward contracts. They do this in several ways. First, in contrast to forward contracts, the quantities delivered and the delivery dates of futures contracts are standardized, making it more likely that different parties can be matched in the futures market, thereby increasing the liquidity of the market. In the case of the 10-year Canada bond contract, the quantity delivered is $100 000 face value of bonds, and the delivery dates are set to be the last business days of March, June, September, and December. Second, after the futures contract has been bought or sold, it can be traded (bought or sold) again at any time up until the delivery date. In contrast, once a forward contract is agreed on, it typically cannot be traded. Third, in a futures contract, not just one specific type of Canada bond is deliverable on the delivery date, as is the case in a forward contract. Instead, any Canada bond that matures in more than 15 years and is not callable for 15 years is eligible for delivery. Allowances for continual trading increase the liquidity of the futures market, as does the option to deliver a range of Canada bonds rather than one specific type of bond.

Another reason why futures contracts specify that more than one bond is eligible for delivery is to limit the possibility that someone might corner the market and "squeeze" traders who have sold contracts. To corner the market, someone buys up all the deliverable securities so that investors with a short position cannot obtain from anyone else the securities that they contractually must deliver on the delivery date. As a result, the person who has cornered the market can set exorbitant prices for the securities that investors with a short position must buy to fulfill their obligations under the futures contract. The person who has cornered the market makes a fortune, but investors with a short position take a terrific loss. Clearly, the possibility that corners might occur in the market will discourage people from taking a short position and might therefore decrease the size of the market. By allowing for the delivery of many different securities, the futures contract makes it harder for any one party to corner the market, because a much larger amount of securities would have to be purchased to establish the corner. Corners are of concern to both regulators and the organized exchanges that design futures contracts.

Trading in the futures market has been organized differently from trading in forward markets to overcome the default risk problems inherent in forward contracts. In both types of markets, for every contract, there must be a buyer who is taking a long position and a seller who is taking a short position. However, the buyer and seller of a futures contract make their contract not with each other but with the clearinghouse associated with the futures exchange. This setup means that, in contrast to the forward market, the buyer of the futures contract does not need to worry about the financial health or trustworthiness of the seller, and vice versa. As long as the clearinghouse is financially solid, buyers and sellers of futures contracts do not have to worry about default risk.

To make sure that the clearinghouse is financially sound and does not run into financial difficulties that might jeopardize its contracts, buyers or sellers of futures contracts must put an initial deposit, called a **margin requirement**, of perhaps $2000 per Canada bond contract into a margin account kept at their brokerage firm. Futures contracts are then **marked to market** every day: At the end of every trading day, the change in the value of the futures contract is added to or subtracted from the margin account. Suppose that after you buy a Canada bond contract at a price of 115 on Wednesday morning, its closing price at the end of the day, the *settlement price*, falls to 114. You now have a loss of 1 point, or $1000, on the contract, and the seller who sold you the contract has a gain of 1 point, or $1000. The $1000 gain is added to the seller's margin account, making a total of $3000 in that account, and the $1000 loss is subtracted from your account so that you now have only $1000 in your account. If the amount in this margin account falls below the maintenance margin requirement (which can be the same as the initial requirement but is usually a little less), the trader is required to add money to the account. For example, if the maintenance margin requirement is also $2000, you will have to add $1000 to your account to bring it up to $2000. Margin requirements and marking to market make it far less likely that a trader will default on a contract, thus protecting the futures exchange from losses.

A final advantage that futures markets have over forward markets is that most futures contracts do not end in delivery of the underlying asset on the expiration date, whereas forward contracts do. A trader who sold a futures contract is allowed to avoid delivery on the expiration date by making an offsetting purchase of a futures contract. Because the simultaneous holding of the long and short positions means that the trader would in effect be delivering the bonds to itself, under the exchange rules the trader is allowed to cancel both contracts. Allowing traders to cancel their contracts in this way lowers the cost of conducting trades in the futures market relative to the forward market: A futures trader can avoid the costs of physical delivery, something that is not so easy to do with forward contracts.

APPLICATION Hedging Foreign Exchange Risk

As we discussed in Chapter 1, foreign exchange rates have been highly volatile in recent years. The large fluctuations in exchange rates subject financial institutions and other businesses to significant foreign exchange risk because such fluctuations generate substantial gains and losses. Luckily for financial institution managers, the financial derivatives discussed in this chapter—forward and financial futures contracts—can be used to hedge foreign exchange risk.

To understand how financial institution managers manage foreign exchange risk, let's suppose that in January, the First National Bank's customer Frivolous Luxuries, Inc., is due a payment of 10 million euros in two months for $15 million worth of goods it has just sold in Germany. Frivolous Luxuries is concerned that if the value of the euro falls substantially from its current value of $1.50, the company might suffer a large loss because the 10 million euro payment will no longer be worth $15 million. So Sam, the CEO of Frivolous Luxuries, calls his friend Mona, the manager of the First National Bank, and asks her to hedge this foreign exchange risk for his company. Let's see how the bank manager could do this by using either forward or financial futures contracts.

Hedging Foreign Exchange Risk with Forward Contracts

Forward markets in foreign exchange have been highly developed by commercial banks and investment banking operations that engage in extensive foreign exchange trading and are widely used to hedge foreign exchange risk. Mona knows that she can use this market to hedge the foreign exchange risk for Frivolous Luxuries. Such a hedge is quite straightforward for her to execute. Because the payment of euros in two months means that Sam will hold a long position in euros at that time, Mona knows that the basic principle of hedging indicates that she should offset this long position by a short position. Thus she simply enters a forward contract that obligates her to sell 10 million euros two months from now in exchange for dollars at the current forward rate of $1.50 per euro.[4]

In two months, when her customer receives the 10 million euros, the forward contract ensures that they will be exchanged for dollars at an exchange rate of $1.50 per euro, thus yielding $15 million. No matter what happens to future exchange rates, Frivolous Luxuries will be guaranteed $15 million for the goods it sold in Germany. Mona calls her friend Sam to let him know that his company is now protected from any foreign exchange movements, and he thanks her for her help.

Hedging Foreign Exchange Risk with Futures Contracts

As an alternative, Mona can use the currency futures market to hedge the foreign exchange risk. In this case, she will do some research and find that the Chicago Mercantile Exchange has a euro contract with a contract amount of 125 000 euros and a price of $1.50 per euro. To conduct the hedge, Mona must sell euros as with the forward contract, to the tune of 10 million euros of the March futures. How many of the Chicago Mercantile Exchange March euro contracts must Mona sell to hedge the 10 million euro payment due in March?

Using Equation 1 with $VA = 10$ million euros and $VC = 125\,000$ euros, we get:

$$NC = 10 \text{ million}/125\,000 = 80.$$

Thus Mona conducts the hedge by selling 80 of the CME euro contracts. Given the $1.50-per-euro price, the sale of the contract yields $80 \times 125\,000$ euros $= \$15$ million. This futures hedge enables her to lock in the exchange rate for Frivolous Luxuries so that it gets its payment of $15 million.

[4]The forward exchange rate will probably differ slightly from the current spot rate of $1 per euro because the interest rates in Germany and Canada probably will not be equal. In that case, the future expected exchange rate will not equal the current spot rate and neither will the forward rate. However, since interest differentials have typically been less than 6% at an annual rate (1% bimonthly), the expected appreciation or depreciation of the euro over a two-month period has always been less than 1%. Thus the forward rate is always close to the current spot rate, and our assumption in the example that the forward rate and the spot rate are the same is a reasonable one.

One advantage of using the futures market is that the contract size of 125 000 euros, worth $187 500, is quite a bit smaller than the minimum size of a forward contract, which is usually $1 million or more. However, in this case, the bank manager is making a large enough transaction that she can use either the forward or the futures market. Her choice depends on whether the transaction costs are lower in one market than in the other. If the First National Bank is active in the forward market, that market would probably have the lower transaction costs. If First National rarely deals in foreign exchange forward contracts, Mona might do better by sticking with the futures market. ◆

13.4 OPTIONS

LO 13.4 Identify the different types of options contracts, and summarize the three conclusions regarding call and put options.

Another vehicle for hedging interest-rate and stock market risk involves the use of *options* on financial instruments. **Options** are contracts that give the purchaser the option, or *right*, to buy or sell the underlying financial instrument at a specified price, called the **exercise price** or **strike price**, within a specific period of time (the *term to expiration*). The seller (sometimes called the *writer*) of the option is *obligated* to buy or sell the financial instrument from or to the purchaser if the owner of the option exercises the right to sell or buy. These option contract features are important enough to be emphasized: The *owner* or buyer of an option does not *have* to exercise the option; they can let the option expire without using it. Hence the owner of an option is not obligated to take any action but rather has the *right* to exercise the contract if they so choose. The seller of an option, by contrast, has no choice in the matter; they *must* buy or sell the financial instrument if the owner exercises the option.

Because the right to buy or sell a financial instrument at a specified price has value, the owner of an option is willing to pay an amount, called a **premium**, for this right. Option contracts are of two types: **American options** can be exercised *at any time up to* the expiration date of the contract, and **European options** can be exercised only *on* the expiration date.

Option contracts are written on a number of financial instruments. Options on individual stocks are called **stock options**, and such options have existed for a long time. Option contracts on financial futures, called **financial futures options** or, more commonly, **futures options**, were developed in 1982 and have become the most widely traded option contracts.

You might wonder why option contracts are more likely to be written on financial futures than on underlying debt instruments such as bonds or certificates of deposit. As we saw earlier in the chapter, on the expiration date, the price of the futures contract and that of the deliverable debt instrument will be the same because of arbitrage. So it would seem that investors should be indifferent about having the option written on the debt instrument or on the futures contract. However, financial futures contracts have been so well designed that their markets are often more liquid than the markets in the underlying debt instruments. So investors would rather have the option contract written on the more liquid instrument—in this case, the futures contract. This explains why the most popular futures options are written on many of the same futures contracts listed in Table 13-1.

In Canada, the regulation of options markets is the responsibility of the Canadian Derivatives Clearing Corporation, a wholly-owned subsidiary of the Montreal Exchange.

The regulation of U.S. options markets is split between the Securities and Exchange Commission, which regulates stock options, and the Commodity Futures Trading Commission, which regulates futures options. Regulation focuses on ensuring that writers of options have enough capital to make good on their contractual obligations and on overseeing traders and exchanges to prevent fraud and to ensure that the market is not being manipulated.

Stock Options

To understand option contracts more fully, let's first discuss stock options (options on individual stocks) before we turn to the more complicated futures options (that is, options on financial futures).

Profits and Losses on Calls

A **call option** is a contract that gives the owner the right (but not the obligation) to *buy* a stock (from the option writer) at the exercise price within a specific period of time. Since a call represents an option to buy, the purchase of a call is undertaken if the price of the underlying stock is expected to go up. The buyer of a call is said to be **long in a call**, and the writer of a call is said to be **short in a call**. The buyer of a call option will have to pay a premium, called a **call premium**, in order to get the writer to sign the contract and assume the risk.

To illustrate the profitability of a call option, suppose you hold a European call option on an equity security with an exercise price of X (say \$100) and a call premium of α (say \$5), as shown in Figure 13-1. If, at the expiration date, the price of the underlying asset, S, is less than X, the call will not be exercised, resulting in a loss

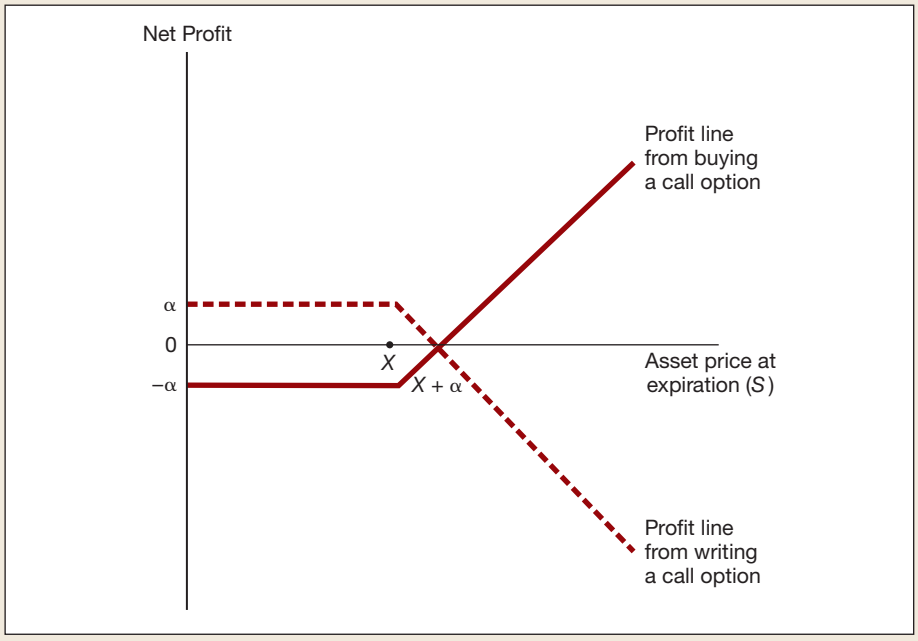


FIGURE 13-1 Profits from Buying and Writing a Call Option

of the premium. At a price above X, the call will be exercised. In particular, at a price between X and $X + \alpha$, the gain will be insufficient to cover the cost of the premium, while at a price above $X + \alpha$, the call will yield a net profit. In fact, at a price above $X + \alpha$, each \$1 rise in the price of the asset will cause the profit of the call option to increase by \$1.

The payoff function from writing the call option is the mirror image of the payoff function from buying the call and is represented by the dashed line in Figure 13-1. Note that the writer of the call receives the call premium up front and must stand ready to sell the underlying asset to the buyer of the call at the exercise price X if the buyer exercises the option to buy.

In general, the value of a call option C at expiration with asset price S (at that time) and exercise price X is

$$C = \max(0, S - X)$$

In other words, the value of a call option at maturity (also known as the intrinsic value) is the difference between the current asset price and the exercise price, $S - X$, or zero, whichever is greater. If the stock price happens to be greater than the exercise price $(S > X)$, the call is said to be in the money, and the owner will exercise it for a net profit of $C - \alpha$. If the asset price happens to be less than the exercise price $(S < X)$, the call is said to be out of the money and will expire worthless. A call with $S = X$ is said to be at the money, or trading at par.

Profit and Losses on Puts As already noted, a second type of option contract is the put option. A **put option** is a contract that gives the owner the right (but not the obligation) to *sell* an asset (to the option writer) at the specified exercise price within a specific period of time. As a put represents an option to sell rather than buy, it is worth buying a put when the price of the underlying asset is expected to fall. As with calls, the buyer of a put is said to be long in a put, and the writer of a put is said to be short in a put. Also, as with calls, the buyer of a put option will have to pay a premium, called a put premium, in order to get the writer to sign the contract and assume the risk.

The solid line in Figure 13-2 illustrates the profit line from buying a put with an exercise price of X and a premium of β. At a price of X or higher, the put will not be exercised, resulting in a loss of the premium. At a price below $X - \beta$, the put is sufficiently deep in the money to cover the option premium and yield a net profit. In fact, between $X - \beta$ and X, the put will be exercised, but the gain is insufficient to cover the cost of the premium.

The payoff function from writing a put is the mirror image of that from buying a put and is represented by the dashed line in Figure 13-2. As with writing a call, the writer of a put receives the put premium β up front and must sell the asset underlying the option if the buyer of the put exercises the option to sell.

In general, the value of a put option P at the expiration date with exercise price X and asset price S (at that time) is

$$P = \max(X - S, 0)$$

That is, the value of a put at maturity is the difference between the exercise price of the option and the price of the asset underlying the option, $X - S$, or zero, whichever is greater. If the stock price happens to be greater than the exercise price $(S > X)$, the put is said to be out of the money and will expire worthless. If the asset price happens to be less than the exercise price $(S < X)$, the put is said to be in the money and the

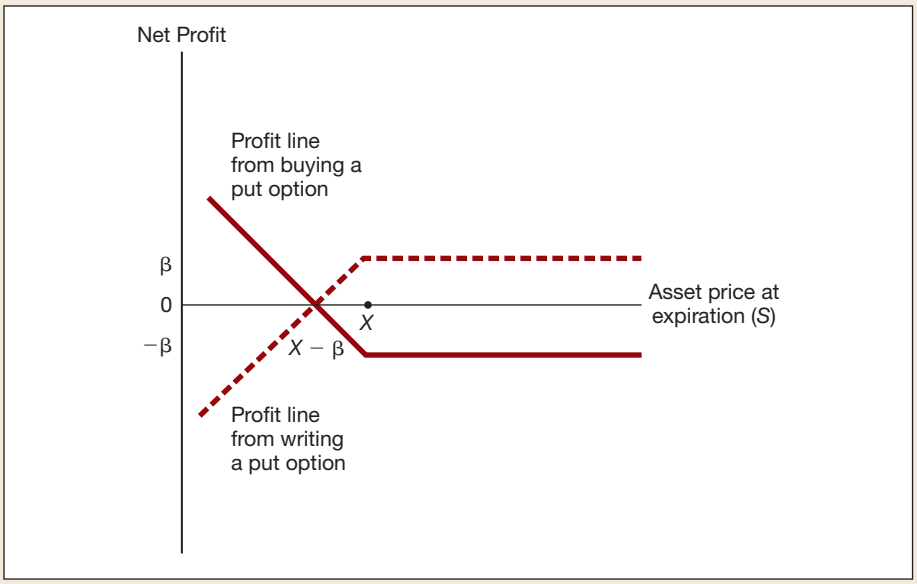


FIGURE 13-2 **Profits from Buying and Writing a Put Option**

owner will exercise it for a net profit of $P - \beta$. If $S = X$, the put is said to be at the money.

Remembering which is a call option and which is a put option is not always easy. To keep them straight, just remember that having a *call* option to *buy* an asset is the same as having the option to *call in* the asset for delivery at a specified price. Having a *put* option to *sell* an asset is the same as having the option to *put up* an asset for the other party to buy.

Futures Options

To understand option contracts on financial futures, let's first examine the option on a June Canada bond futures contract. If you bought this futures contract at a price of 115 (that is, $115 000), you agreed to pay $115 000 for $100 000 face value of long-term Canada bonds when they are delivered to you at the end of June. If you sold this futures contract at a price of 115, you agreed, in exchange for $115 000, to deliver $100 000 face value of the long-term Canada bonds at the end of June. An option contract on the Canada bond futures contract has several key features: (1) It has the same expiration date as the underlying futures contract; (2) it is an American option and so can be exercised at any time before the expiration date; and (3) the premium (price) of the option is quoted in points that mimic the futures contract, so each point corresponds to $1000. If, for a premium of $2000, you buy one call option contract on the June Canada bond contract with an exercise price of 115, you have purchased the right to buy (call in) the June Canada bond futures contract for a price of 115 ($115 000 per contract) at any time until the expiration date of this contract at the end of June. Similarly, if you buy for $2000 a put option on the June Canada bond contract with an

exercise price of 115, you have the right to sell (put up) the June Canada bond futures contract for a price of 115 ($115 000 per contract) at any time until the end of June.

Futures option contracts are somewhat complicated, so to explore how they work and how they can be used to hedge risk, let's first examine how profits and losses occur on the call option on the June Canada bond futures contract. In February, our old friend Irving the Investor buys, for a $2000 premium, a call option on the $100 000 June Canada bond futures contract with a strike price of 115. (We will assume that if Irving exercises the option, he does so on the expiration date at the end of June and not before.) Suppose that the underlying Canada bond for the futures contract has a price of 110 on the expiration date at the end of June. Recall that on the expiration date, arbitrage forces the price of the futures contract to be the same as the price of the underlying bond, so the contract also has a price of 110 on the expiration date at the end of June. If Irving exercises the call option and buys the futures contract at an exercise price of 115, he will lose money by buying at 115 and selling at the lower market price of 110. Because Irving is smart, he will not exercise the option, but he will be out the $2000 premium he paid. In such a situation, in which the price of the underlying financial instrument is below the exercise price, a call option is said to be "out of the money." At the price of 110 (less than the exercise price), Irving thus suffers a loss on the option contract of the $2000 premium he paid. This loss is plotted as point A in panel (a) of Figure 13-3.

On the expiration date, if the price of the futures contract is 115, the call option is "at the money," and Irving is indifferent regarding his option to buy the futures contract, because exercising the option at 115 when the market price is also at 115 produces no gain or loss. Because he has paid the $2000 premium, at the price of 115 his contract again has a net loss of $2000, plotted as point B. If the futures contract instead has a price of 120 on the expiration day, the option is "in the money," and Irving benefits from exercising the option: He will buy the futures contract at the exercise price of 115 and then sell it for 120, thereby earning a 5-point gain ($5000 profit) on the $100 000 Canada bond contract. Because Irving paid a $2000 premium for the option contract, however, his net profit is $3000 ($5000 − $2000). The $3000 profit at a price of 120 is plotted as point C. Similarly, if the price of the futures contract rises to 125, the option contract will yield a net profit of $8000 ($10 000 from exercising the option minus the $2000 premium), plotted as point D. Connecting these points, we get the kinked profit curve for the call option that we see in panel (a).

Suppose that instead of purchasing the futures *option* contract in February, Irving decides instead to buy the $100 000 June Canada bond *futures* contract at the price of 115. If the price of the bond on the expiration day at the end of June declines to 110, meaning that the price of the futures contract also falls to 110, Irving suffers a loss of 5 points, or $5000. The loss of $5000 on the futures contract at a price of 110 is plotted as point A in panel (a). At a price of 115 on the expiration date, Irving will have a zero profit on the futures contract, plotted as point B′. At a price of 120, Irving will have a profit of 5 points on the contract, or $5000 (point C′), and at a price of 125, the profit will be 10 percentage points, or $10 000 (point D′). Connecting these points, we get the linear (straight-line) profit curve for the futures contract, graphed in panel (a).

Now we can see the major difference between a futures contract and an option contract. As indicated by the profit curve for the futures contract in panel (a), the futures contract has a linear profit function: Profits grow by an equal dollar amount for every point increase in the price of the underlying financial instrument. By contrast, the kinked profit curve for the option contract is nonlinear, meaning that profits do not always grow by an equal dollar amount with a given change in the price of the underlying financial instrument. This nonlinearity indicates that the call option protects Irving

Mini-lecture

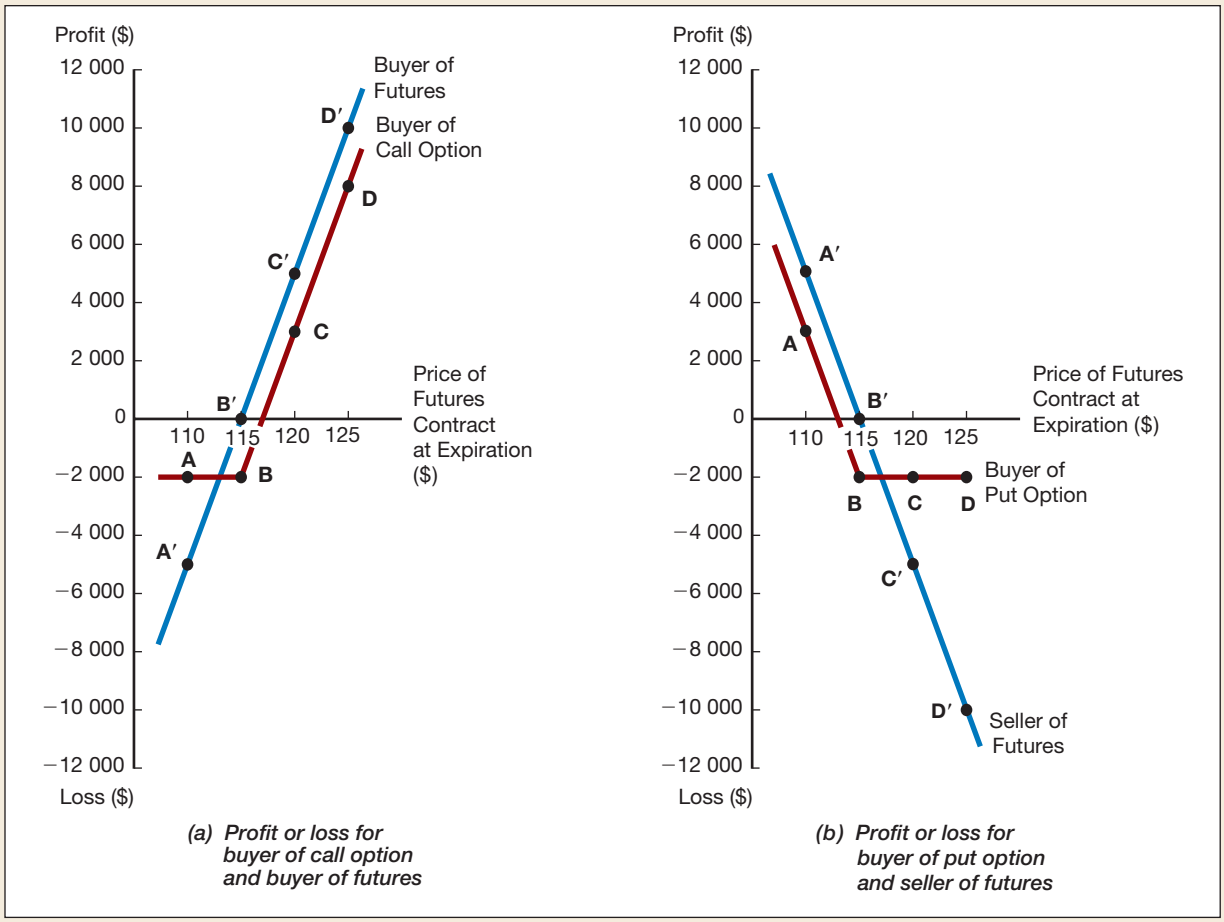


FIGURE 13-3 Profits and Losses on Options Versus Futures Contracts

The futures contract is the $100 000 June Canada bond contract, and the option contracts are written on this futures contract with an exercise price of 115. Panel (a) shows the profits and losses for the buyer of the call option and the buyer of the futures contract, and panel (b) shows the profits and losses for the buyer of the put option and the seller of the futures contract.

from suffering losses that are greater than the amount of the $2000 premium. In contrast, Irving's loss on the futures contract is $5000 if the price on the expiration day falls to 110, and if the price falls even further, Irving's loss will be even greater.

This insurance-like feature of option contracts explains why their purchase price is referred to as a premium. Once the underlying financial instrument's price rises above the exercise price, however, Irving's profits will grow linearly. Irving has given up something by buying an option rather than a futures contract. As we see in panel (a), when the price of the underlying financial instrument rises above the exercise price, Irving's profits are always less than the profits he would have made on the futures contract—by exactly the $2000 premium he paid.

Panel (b) plots the results of the same profit calculations if Irving buys not a call option but a put option (an option to sell) with an exercise price of 115 for a premium

of $2000, and if he sells the futures contract rather than buying it. In this case, if on the expiration date the Canada bond futures have a price above the 115 exercise price, the put option is "out of the money." Irving will not want to exercise the put option because by doing so he would have to sell the futures contract he owns at a price below the market price, and he would lose money. He therefore will not exercise his option, and he will be out only the $2000 premium he paid. Once the price of the futures contract falls below the 115 exercise price, Irving benefits from exercising the put option because he can sell the futures contract at a price of 115 but can buy it at a price less than 115. In such a situation, in which the price of the underlying instrument is below the exercise price, the put option is "in the money," and profits rise linearly as the price of the futures contract falls. The profit function for the put option illustrated in panel (b) of Figure 13-3 is kinked, indicating that Irving is protected from losses greater than the amount of the premium he paid. The profit curve for the sale of the futures contract is just the negative of the profit curve for the futures contract in panel (a) and is therefore linear.

Panel (b) of Figure 13-3 confirms the conclusion from panel (a) that profits on option contracts are nonlinear, but profits on futures contracts are linear.

Two additional differences between futures and option contracts must be mentioned. The first is that the initial investment on the two types of contracts differs. As we saw earlier in the chapter, when a futures contract is purchased, the investor must put a fixed amount, called the margin requirement, into a margin account. But when an option contract is purchased, the initial investment is the premium that is paid for the contract. The second important difference between the contracts is that with a futures contract, money changes hands daily when the contract is marked to market, whereas with an option contract, money changes hands only when the option is exercised.

APPLICATION Hedging with Futures Options

Earlier in the chapter, we saw how the First National Bank could hedge the interest-rate risk on its $5 million holdings of 6s of 2041 by selling $5 million of Canada bond futures. A rise in interest rates and the resulting fall in bond prices and bond futures contracts would lead to profits on the bank's sale of the futures contracts that would exactly offset the losses on the 6s of 2041 that the bank is holding.

As suggested in panel (b) of Figure 13-3, an alternative way for the manager to protect against a rise in interest rates and hence a decline in bond prices is to buy $5 million of put options written on the same Canada bond futures. As long as the exercise price is not too far from the current price, as in panel (b), the rise in interest rates and decline in bond prices will lead to profits on the futures and the futures put options, profits that will offset any losses on the $5 million of Canada bonds.

The one problem with using options rather than futures is that the First National Bank will have to pay premiums on the options contracts, thereby lowering the bank's profits in an effort to hedge the interest-rate risk. Why might the bank manager be willing to use options rather than futures to conduct the hedge? The answer is that the option contract, unlike the futures contract, allows the First National Bank to gain if interest rates decline and bond prices rise. With the hedge using futures contracts, the First National Bank does not gain from increases in bond prices because the profits on the bonds it is holding are offset by the losses from the futures contracts it has sold. However, as indicated in panel (b) of Figure 13-3, the situation when the hedge is

conducted with put options is quite different: Once bond prices rise above the exercise price, the bank does not suffer additional losses on the option contracts. At the same time, the value of the Canada bonds the bank is holding will increase, thereby leading to a profit for the bank. Thus by using options rather than futures to conduct the micro hedge, the bank can protect itself from rises in interest rates but will still benefit from interest-rate declines (although the bank's profit will be reduced by the amount of the premium).

Similar reasoning indicates that the bank manager might prefer to use options to conduct the macro hedge to immunize the entire bank portfolio from interest-rate risk. Again, by using options rather than futures, the First National Bank will have to pay the premiums on these contracts up front. By contrast, the use of options allows the bank to keep the gains from a decline in interest rates (which will raise the value of the bank's assets relative to its liabilities), because these gains will not be offset by large losses on the option contracts.

In the case of a macro hedge, there is another reason why the bank might prefer option contracts to futures contracts. Profits and losses on futures contracts can cause accounting problems for banks because such profits and losses are not allowed to be offset by changes in the value of the rest of the bank's portfolio that are not yet realized by the actual sale of the assets. Consider what happens if interest rates fall. If First National sells futures contracts to conduct the macro hedge, then if interest rates fall and the prices of the Canada bond futures contracts rise, the bank will suffer large losses on these contracts. Of course, these losses will be offset by profits in the rest of the bank's portfolio that are not yet realized by actual sale of the assets, but the bank is not allowed to offset these losses in its accounting statements. So even though the macro hedge is serving its intended purpose of immunizing the bank's portfolio from interest-rate risk, the bank will experience large accounting losses if interest rates fall. Indeed, bank managers have lost their jobs when perfectly sound hedges with interest-rate futures generated large accounting losses. Not surprisingly, for this reason bank managers might shrink from using financial futures to conduct macro hedges.

Futures options, however, can come to the rescue of the managers of banks and other financial institutions. Suppose that First National conducts the macro hedge by buying put options rather than selling Canada bond futures. Now, if interest rates fall and bond prices rise well above the exercise price, the bank will not suffer large losses on the option contracts because it will just decide not to exercise its options. The bank will not suffer the accounting problems that it would have suffered if it had hedged with financial futures. Because of the accounting advantages of using futures options to conduct macro hedges, option contracts have become important tools for financial institution managers who are looking to hedge interest-rate risk. ◆

Factors Affecting Option Premiums

Several interesting facts can be noted about how the premiums on option contracts are priced. First, when the strike (exercise) price for a contract is set at a higher level, the premium for the call option is lower and the premium for the put option is higher. For example, in going from a contract with a strike price of 112 to one with a strike price of 115, the premium for a call option for the month of March might fall from $1\frac{45}{64}$ to $\frac{16}{64}$, and the premium for the March put option might rise from $\frac{19}{64}$ to $1\frac{54}{64}$.

Our understanding of the profit function for option contracts, illustrated in Figure 13-3, helps explain this fact. As we saw in panel (a), a lower price for the underlying financial

instrument (in this case, a Canada bond futures contract) relative to the option's exercise price results in lower profits on the call (buy) option. Thus, the higher the strike price, the lower the profits on the call option contract and the lower the premium that investors like Irving will be willing to pay. Similarly, we saw in panel (b) that a lower price for the underlying financial instrument relative to the exercise price increases profits on the put (sell) option so that a higher strike price increases profits and thus causes the premium to increase.

Second, as the period of time over which the option can be exercised (the term to expiration) increases, the premiums for both call and put options rise. For example, at a strike price of 112, the premium on a call option might increase from $1\frac{45}{64}$ in March to $1\frac{50}{64}$ in April to $2\frac{28}{64}$ in May. Similarly, the premium on a put option might increase from $\frac{19}{64}$ in March to $1\frac{43}{64}$ in April to $2\frac{22}{64}$ in May. The fact that the premium increases with increases in the term to expiration is also explained by the nonlinear profit function for option contracts. As the term to expiration lengthens, there is a greater chance that the price of the underlying financial instrument will be very high or very low on the expiration date. If the price becomes very high and goes well above the exercise price, the call (buy) option will yield a high profit; if the price becomes very low and goes well below the exercise price, the losses will be small because the owner of the call option will simply decide not to exercise the option. The possibility of greater variability of the price of the underlying financial instrument as the term to expiration lengthens raises profits, on average, for the call option.

Similar reasoning tells us that the put (sell) option also becomes more valuable as the term to expiration increases, because the possibility of greater price variability of the underlying financial instrument increases as the term to expiration increases. The greater chance of a low price increases the chance that profits on the put option will be very high. But the greater chance of a high price does not increase the chance of substantial losses on the put option, because the owner again will just decide not to exercise the option.

We can think about this feature of option contracts in another way: Option contracts have an element of "heads, I win; tails, I don't lose too badly." The greater variability of the final price on the expiration date increases the value of both kinds of options. Because a longer term to the expiration date leads to greater variability in the final price on the expiration date, a longer term to expiration raises the value of an option contract.

The reasoning we just developed explains another important fact about option premiums. When the volatility of the price of the underlying instrument is great, the premiums on both call and put options are higher. Higher volatility of price means that, for a given expiration date, there will again be greater variability in the final price on the expiration date. The "heads, I win; tails, I don't lose too badly" property of options implies that the greater variability of possible prices on the expiration date increases average profits for the option and thus increases the premium that investors are willing to pay.

Summary

Our analysis in this application leads us to the following conclusions about the factors that determine the premium on an option contract:

1. The higher the strike price, everything else being equal, the lower the premium on call (buy) options and the higher the premium on put (sell) options.
2. The greater the term to expiration, everything else being equal, the higher the premiums on both call and put options.
3. The greater the volatility of prices of the underlying financial instrument, everything else being equal, the higher the premiums on both call and put options.

The results we have derived here appear in more formal models, such as the Black-Scholes model, that analyze how the premiums on options are priced. You might study such models in finance courses.

13.5 SWAPS

LO 13.5 Define a swap, and summarize the advantages and disadvantages of interest-rate swaps.

In addition to forwards, futures, and options, financial institutions use one other important financial derivative to manage risk. **Swaps** are financial contracts that obligate each party to the contract to exchange (swap) a set of payments (not assets) that it owns for another set of payments that are owned by another party. Swaps are of two basic kinds. **Currency swaps** involve the exchange of a set of payments in one currency for a set of payments in another currency. **Interest-rate swaps** involve the exchange of one set of interest payments for another set of interest payments, all denominated in the same currency.

Interest-Rate Swap Contracts

Interest-rate swaps are an important tool for managing interest-rate risk. They first appeared in the United States in 1982, when, as we have seen, demand increased for financial instruments that could be used to reduce interest-rate risk. The most common type of interest-rate swap (called the *plain vanilla swap*) specifies (1) the interest rate on the payments that are being exchanged; (2) the type of interest payments (variable or fixed-rate); (3) the amount of **notional principal**, which is the amount on which the interest is being paid; and (4) the time period over which the exchanges will continue to be made. There are many other, more complicated versions of swaps, including forward swaps and swap options (called *swaptions*), but here we will look only at the plain vanilla swap. Figure 13-4 illustrates an interest-rate swap between the Midwest Savings Bank and

Mini-lecture

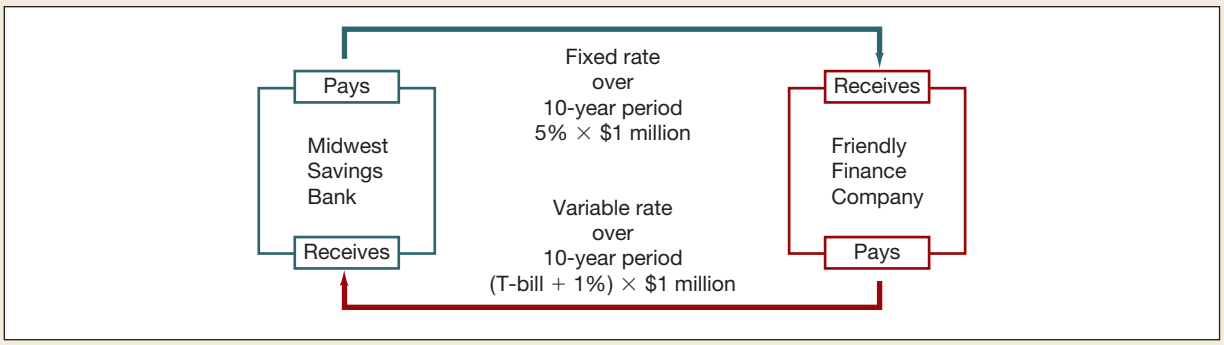


FIGURE 13-4 Interest-Rate Swap Payments

In this swap arrangement, with a notional principal of $1 million and a term of 10 years, the Midwest Savings Bank agrees to pay a fixed rate of 5% × $1 million to the Friendly Finance Company, which in turn agrees to pay the one-year Treasury bill rate plus 1% × $1 million to the Midwest Savings Bank.

the Friendly Finance Company. Midwest Savings agrees to pay Friendly Finance a fixed rate of 5% on $1 million of notional principal for the next 10 years, and Friendly Finance agrees to pay Midwest Savings the one-year Treasury bill rate plus 1% on $1 million of notional principal for the same period. Thus, as shown in Figure 13-4, every year the Midwest Savings Bank will pay the Friendly Finance Company 5% on $1 million, while Friendly Finance will pay Midwest Savings the one-year T-bill rate plus 1% on $1 million.

APPLICATION Hedging with Interest-Rate Swaps

You might wonder why these two parties would find it advantageous to enter into this swap agreement. The answer is that it may help both of them hedge interest-rate risk.

Suppose that Midwest Savings Bank, which tends to borrow short-term and then lend long-term in the mortgage market, has $1 million less of rate-sensitive assets than it has of rate-sensitive liabilities. As we learned in Chapter 12, this situation means that as interest rates rise, the increase in the cost of funds (liabilities) will be greater than the increase in interest payments the bank receives on its assets, many of which are fixed-rate. Rising interest rates thus will lead to a shrinking of Midwest Savings' net interest margin and a decline in its profitability. As we saw in Chapter 12, to avoid this interest-rate risk, Midwest Savings would like to convert $1 million of its fixed-rate assets into $1 million of rate-sensitive assets, in effect making rate-sensitive assets equal to rate-sensitive liabilities, thereby eliminating the gap. This is exactly what happens when Midwest Savings engages in the interest-rate swap. By taking $1 million of its fixed-rate income and exchanging it for $1 million of rate-sensitive Treasury bill income, it has converted income on $1 million of fixed-rate assets into income on $1 million of rate-sensitive assets. Now, when interest rates rise, the increase in rate-sensitive income on the bank's assets will exactly match the increase in the rate-sensitive cost of funds on its liabilities, leaving the net interest margin and bank profitability unchanged.

The Friendly Finance Company, which issues long-term bonds to raise funds and uses them to make short-term loans, finds itself in exactly the opposite situation of Midwest Savings: It has $1 million more of rate-sensitive assets than of rate-sensitive liabilities. It is therefore concerned that a fall in interest rates, which will result in a larger drop in income from its assets than the decline in the cost of funds on its liabilities, will cause a decline in profits. By conducting the interest-rate swap, Friendly Finance eliminates this interest-rate risk because it has converted $1 million of rate-sensitive income into $1 million of fixed-rate income. Now the Friendly Finance Company finds that when interest rates fall, the decline in rate-sensitive income is smaller and so is matched by the decline in the rate-sensitive cost of funds on its liabilities, leaving its profitability unchanged.

Advantages of Interest-Rate Swaps

To eliminate interest-rate risk, both the Midwest Savings Bank and the Friendly Finance Company could have rearranged their balance sheets by converting fixed-rate assets into rate-sensitive assets, and vice versa, instead of engaging in an interest-rate swap. However, this strategy would have been costly for both financial institutions for

several reasons. First, financial institutions incur substantial transaction costs when they rearrange their balance sheets. Second, different financial institutions have different informational advantages and often make loans to select customers who prefer certain maturities. Thus, by adjusting the balance sheet to eliminate interest-rate risk, the institution might lose these informational advantages, which most financial institutions are unwilling to give up. Interest-rate swaps solve these problems for financial institutions because, in effect, they allow the institutions to convert fixed-rate assets into rate-sensitive assets without affecting the balance sheet. Large transaction costs are avoided, and the financial institutions can continue to use their informational advantage to make loans to preferred customers.

We have learned that financial institutions can also hedge interest-rate risk with other financial derivatives, such as futures contracts and futures options. Interest-rate swaps have one big advantage over these other derivatives: They can be written for very long horizons, sometimes as long as 20 years, whereas financial futures and futures options typically have much shorter horizons, usually not much more than a year. If a financial institution needs to hedge interest-rate risk for a long horizon, financial futures and option markets may not do it much good. Instead, it can turn to the swap market.

Disadvantages of Interest-Rate Swaps

Although interest-rate swaps have important advantages that make them very popular with financial institutions, they also have weaknesses that limit their usefulness. Swap markets, like forward markets, can suffer from a lack of liquidity. Let's return to the swap between the Midwest Savings Bank and the Friendly Finance Company. As is sometimes the case with a forward contract, it might be difficult for the Midwest Savings Bank to link up with the Friendly Finance Company to arrange the swap. In addition, even if the Midwest Savings Bank is able to find a willing counterparty like the Friendly Finance Company, it might not be able to negotiate a good deal because it has no other institution with which to negotiate.

Swap contracts also are subject to the same default risk as forward contracts. If interest rates rise, the Friendly Finance Company would love to get out of the swap contract, because the fixed-rate interest payments it receives are less than the rates it could get in the open market. It might then default on the contract, exposing Midwest Savings to a loss. Alternatively, the Friendly Finance Company could go bust, meaning that the terms of the swap contract would not be fulfilled.

Financial Intermediaries in Interest-Rate Swaps

As we have just seen, financial institutions do have to be aware of the possibility of losses from a default on swaps. As with a forward contract, each party to a swap must have a lot of information about the other party to make sure that the contract is likely to be fulfilled. The need for information about counterparties and the liquidity problems in swap markets can limit the usefulness of these markets. However, as we saw in Chapter 8, when informational and liquidity problems crop up in a market, financial intermediaries come to the rescue. That is exactly what happens in swap markets. Intermediaries such as investment banks and especially large commercial banks are able to acquire information cheaply about the creditworthiness and reliability of parties to swap contracts and are also able to match parties to a swap. Hence large commercial banks and investment banks have set up swap markets in which they act as intermediaries. ◆

13.6 CREDIT DERIVATIVES

LO 13.6 Summarize the three types of credit derivatives.

In recent years, a new type of derivative to hedge credit risk has arrived on the scene. Like other derivatives, **credit derivatives** offer payoffs on previously issued securities, but only on securities that bear credit risk. In the past 10 years, the markets in credit derivatives have grown at an astounding pace, and the notional amounts of these derivatives now number in the trillions of dollars. These credit derivatives take several forms.

Credit Options

Credit options work just like the options discussed earlier in the chapter: For a fee, the purchaser gains the right to receive profits that are tied either to the price of an underlying security or to an interest rate. Suppose you buy $1 million of General Motors bonds but worry that a potential slowdown in the sale of SUVs might lead a credit-rating agency to *downgrade* (lower the credit rating on) GM bonds. As we saw in Chapter 6, such a downgrade would cause the price of GM bonds to fall. To protect yourself, you could buy an option—for, say, $15 000—to sell the $1 million of bonds at a strike price that is the same as the current price. With this strategy, you would not suffer any losses if the value of the GM bonds declined because you could exercise the option and sell them at the price you paid for them. In addition, you would be able to reap any gains if GM bonds rose in value.

The second type of credit option ties profits to changes in an interest rate, such as a credit spread (the interest rate on the average bond with a particular credit rating minus the interest rate on default-free bonds, such as those issued by the Canadian government). Suppose that your company, which has a Baa credit rating, plans to issue $10 million of one-year bonds in three months and expects to have a credit spread of one percentage point (i.e., it will pay an interest rate that is one percentage point higher than the one-year T-bill rate). You are concerned that the market might start to think that Baa companies in general will become riskier in the coming months. If this were to happen by the time you are ready to issue your bonds in three months, you would have to pay a higher interest rate than the one percentage point in excess of the T-bill rate, and your cost of issuing the bonds would increase. To protect yourself against these higher costs, you could buy—for, say, $20 000—a credit option on $10 million of Baa bonds that would pay you the difference between the average Baa credit spread in the market and the one percentage point credit spread on $10 million. If the credit spread jumps to two percentage points, you would receive $100 000 from the option $(=[2\% - 1\%] \times \$10 \text{ million})$, which would exactly offset the $100 000 higher interest costs from the one percentage point higher interest rate you would have to pay on your $10 million of bonds.

Credit Swaps

Suppose you manage a bank in Calgary called Oil Drillers' Bank (ODB), which specializes in lending to a particular industry in your local area, oil drilling companies. Another bank, Potato Farmers Bank (PFB), specializes in lending to potato farmers in Charlottetown. Both ODB and PFB have a problem because their loan portfolios are not sufficiently diversified. To protect ODB against a collapse in the oil market, which would result in defaults on most of its loans made to oil drillers, you could reach

an agreement to have the loan payments on, say, $100 million worth of your loans to oil drillers paid to the PFB in exchange for PFB paying you the loan payments on $100 million of its loans to potato farmers. Such a transaction, in which risky payments on loans are swapped for each other, is called a **credit swap**. As a result of this swap, ODB and PFB have increased their diversification and lowered the overall risk of their loan portfolios because some of the loan payments to each bank are now coming from a different type of loan.

Another form of credit swap is, for arcane reasons, called a **credit default swap**, although it functions more like insurance. With a credit default swap, one party that wants to hedge credit risk pays a fixed payment on a regular basis in return for a contingent payment that is triggered by a *credit event,* such as the bankruptcy of a particular firm or the downgrading of the firm's credit rating by a credit-rating agency. For example, you could use a credit default swap to hedge the $1 million of General Motors bonds that you are holding by arranging to pay an annual fee of $1000 in exchange for a payment of $10 000 if the GM bonds' credit rating is lowered. If a credit event happens and GM's bonds are downgraded so that their price falls, you will receive a payment that will offset some of the loss you will suffer if you sell the bonds at this lower price.

Credit-Linked Notes

Another type of credit derivative, the **credit-linked note**, is a combination of a bond and a credit option. Just like any corporate bond, the credit-linked note makes periodic coupon (interest) payments and a final payment of the face value of the bond at maturity. If a key financial variable specified in the note changes, however, the issuer of the note has the right (option) to lower the payments on the note. For example, General Motors could issue a credit-linked note that pays a 5% coupon rate with the specification that if a national index of SUV sales falls by 10%, then GM has the right to lower the coupon rate by two percentage points, to 3%. In this way, GM can lower its risk because when it is losing money as SUV sales fall, it can offset some of these losses by making smaller payments on its credit-linked notes.

APPLICATION Lessons from the Global Financial Crisis: When Are Financial Derivatives Likely to Be a Worldwide Ticking Time Bomb?

Although financial derivatives can be useful in hedging risk, the AIG blowup discussed in Chapter 11 illustrates that they can pose a real danger to the financial system. Indeed, Warren Buffet warned about the dangers of financial derivatives by characterizing them as "financial weapons of mass destruction." Particularly scary are the notional amounts of derivatives contracts—around $500 trillion worldwide. How can we tell when financial derivatives are likely to be a ticking time bomb that might explode at any minute, bringing down the world financial system? We can find some answers by studying the recent global financial crisis.

Two major concerns surround financial derivatives. The first is that financial derivatives allow financial institutions to increase their leverage; that is, these institutions can, in effect, hold an amount of the underlying asset that is many times greater

than the amount of money they had to put up. Increasing their leverage enables them to take huge bets, which, if the bets turn sour, can bring down the institution. This is exactly what AIG did, to its great regret, when it plunged into the credit default swap market. Even more of a problem was the fact that AIG's speculation in the credit default swap market had the potential to bring down the whole global financial system. An important lesson from the global financial crisis is that allowing one player to take huge positions in a derivatives market is a highly dangerous move.

A second concern is that banks often hold huge notional amounts of financial derivatives—in particular, interest-rate and currency swaps—that greatly exceed the amount of bank capital, and so these derivatives might expose the banks to serious risk of failure. Banks are indeed major players in the financial derivatives markets, particularly in the interest-rate and currency swaps markets. Our earlier analysis showed that banks are the natural market makers because they can act as intermediaries between two counterparties that would not necessarily be able to make a swap without the bank's involvement. However, looking at the notional amount of interest-rate and currency swaps at banks gives us a very misleading picture of the banks' risk exposure. Because banks act as intermediaries in the swap markets, they typically are exposed only to credit risk—the risk of default by one of their counterparties. Furthermore, these swaps, unlike loans, do not involve payments of the notional amount but rather much smaller payments that are *based on* the notional amount. For example, in the case of a 7% interest rate, the payment is only $70 000 for a $1 million swap. Estimates of the credit risk exposure from swap contracts indicate that they are on the order of only 1% of the notional value of the contracts and that credit risk exposure at banks from derivatives is generally less than a quarter of their total credit risk exposure from loans. Banks' credit risk exposures from their derivative positions are thus not out of line with other credit risk exposures they face. Indeed, during the recent global financial crisis, in which the entire financial system was put under great stress, derivatives exposure at banks was not a serious problem.

In conclusion, recent events indicate that financial derivatives pose serious dangers to the financial system, but some of these dangers have been overplayed. The biggest danger lies in the trading activities of financial institutions, particularly their trading in credit derivatives, as was illustrated by AIG's activities in the credit default swap market. As discussed in Chapter 9, regulators have been paying increased attention to this danger and are continuing to develop new disclosure requirements and regulatory guidelines that outline rules for derivatives trading. Financial institutions need to disclose their risk exposure in derivatives contracts so that regulators can make sure that a large institution is not playing too big a role in the derivatives market and does not carry too large an exposure to derivatives relative to its capital, as was the case with AIG. Another concern is that derivatives, particularly credit derivatives, must have better clearing mechanisms, that is, the methods of settling contracts, so that the failure of one institution does not bring down many other institutions whose net derivatives positions are small even though they have many offsetting positions. Better clearing could be achieved either by trading these derivatives in an organized exchange, like a futures market, or by having one clearing organization net out trades.

The credit risk exposure posed by interest-rate derivatives, in contrast, seems to be manageable using the standard methods of dealing with credit risk. New regulations for derivatives markets are sure to come in the wake of the global financial crisis. The industry has been given a wake-up call regarding the potential dangers of derivatives products. Now the hope is that any time bomb arising from derivatives can be defused with appropriate efforts on the parts of markets and regulators. ◆

SUMMARY

1. Interest-rate forward contracts, which are agreements to sell a debt instrument at a future (forward) point in time, can be used to hedge (protect against) interest-rate risk. The advantage of forward contracts is that they are flexible; the disadvantages are that they are subject to default risk and their market is illiquid.
2. A financial futures contract is similar to an interest-rate forward contract in that it specifies that a debt instrument must be delivered by one party to another on a stated future date. However, it has advantages over a forward contract in that it is not subject to default risk and is more liquid. Forward and futures contracts can be used by financial institutions to hedge interest-rate risk.
3. An option contract gives the purchaser the right to buy (call option) or sell (put option) a security at the exercise (strike) price within a specific period of time. The profit function for options is nonlinear—profits do not always grow by an equal dollar amount with a given change in the price of the underlying financial instrument. The nonlinear profit function for options explains why their value (as reflected by the premium paid for them) is negatively related to the exercise price for call options, positively correlated to the exercise price for put options, positively related to the term to expiration for both call and put options, and positively related to the volatility of the price of the underlying financial instrument for both call and put options. Financial institutions use futures options to hedge interest-rate risk in a similar fashion to the way they use financial futures and forward contracts. Futures options may be preferred over futures contracts when conducting macro hedges because futures options lead to fewer accounting problems.
4. Interest-rate swaps involve the exchange of one set of interest payments for another set of interest payments. They have default risk and liquidity problems similar to those of forward contracts. As a result, interest-rate swaps often involve intermediaries such as large commercial banks and investment banks that make a market in swaps. Financial institutions find that interest-rate swaps are useful ways to hedge interest-rate risk. Interest-rate swaps have one big advantage over financial futures and options: They can be written for very long horizons.
5. Credit derivatives are a new type of derivative that offers payoffs on previously issued securities that have credit risk. These derivatives—credit options, credit swaps, and credit-linked notes—can be used to hedge credit risk.
6. There are two major concerns about derivatives: They allow financial institutions to more easily increase their leverage and take big bets (by effectively enabling them to hold a larger amount of the underlying assets than the amount of money put down), and they expose financial institutions to large credit risks because the huge notional amounts of derivative contracts greatly exceed the capital of these institutions. The second danger seems to have been overplayed, but the danger from increased leverage using derivatives is very real, as was revealed during the global financial crisis.

MyLab Economics Make the grade with MyLab Economics: Select end-of-chapter questions can be found on MyLab Economics. You can practise them as often as you want, and most feature step-by-step guided instructions to help you find the right answer.

QUESTIONS

1. What are the advantages and disadvantages of using forward contracts to hedge?
2. What advantages do futures contracts have over forward contracts?
3. Suppose that the Bank of Canada did not increase the overnight interest rate target but somehow let the markets know that there could be a rise in the near future. How do you think that financial managers will react to

this news? Which instruments can they use to hedge against a change in interest rates?

4. What are the advantages and disadvantages of using an options contract rather than a futures contract?
5. Explain why greater volatility or a longer term to maturity leads to a higher premium on both call and put options.
6. Why does a lower strike price imply that a call option will have a higher premium and a put option a lower premium?
7. What are the advantages and disadvantages of using interest-rate swaps?
8. If the finance company you manage has a gap of + $5 million (rate-sensitive assets greater than rate-sensitive liabilities by $5 million), describe an interest-rate swap that would eliminate the company's gap.
9. If the bank you manage has a gap of −$42 million, describe an interest-rate swap that would eliminate the bank's income risk from changes in interest rates.
10. How can financial derivatives create excessive risk in the financial system?

APPLIED PROBLEMS

11. If the pension fund you manage expects to have an inflow of $120 million six months from now, what forward contract would you seek to enter into to lock in current interest rates?
12. If the portfolio you manage is holding $25 million of 6s of 2041 Canada bonds with a price of 110, what forward contract would you enter into to hedge the interest-rate risk on these bonds over the coming year?
13. Suppose you buy a call option on a $100 000 Canada bond futures contract with an exercise price of 105. If the price of the Canada bond is 115 at expiration, is the option at the money, in the money, or out of the money? Determine the premium if the profit equals $8000.
14. Jason bought a put option on a $100 000 Canada bond futures contract with an exercise price of 105 for a premium of $2000. If on expiration the futures contract sells for 110, determine Jason's profits and explain if he will exercise the right to sell the futures contract. How would your answer change if the futures contract sells for 95 on expiration?
15. If your company has a payment of 200 million euros due one year from now, how would you hedge the foreign exchange risk in this payment with 125 000-euro futures contracts?
16. If your company has to make a 10 million euro payment to a German company three months from now, how would you hedge the foreign exchange risk in this payment with a 125 000-euro futures contract?
17. Suppose your company will be receiving 30 million euros six months from now, and the euro is currently selling for 1 euro per dollar. If you want to hedge the foreign exchange risk in this payment, what kind of forward contract should you enter into?
18. Suppose the pension fund you are managing is expecting an inflow of funds of $100 million next year, and you want to make sure that you will earn the current interest rate of 8% when you invest the incoming funds in long-term bonds. How would you use the futures market to do this?
19. If you buy a put option on a $100 000 Canada bond futures contract with an exercise price of 95 and the price of the Canada bond is 120 at expiration, is the contract in the money, out of the money, or at the money? What is your profit or loss on the contract if the premium was $4000?
20. Suppose that you buy a call option on a $100 000 Canada bond futures contract with an exercise price of 110 for a premium of $1500. If, upon expiration, the futures contract has a price of 111, what is your profit or loss on the contract?
21. Canada Bank and Trust has $100 million in assets and $83 million in liabilities. The duration of the assets is 5.9 years, and the duration of the liabilities is 1.8 years. How many futures contracts does this bank need to fully hedge itself against interest-rate risk? The available Canada bond futures contracts have a duration of 10 years, have a face value of $1 000 000, and are selling for $979 000.
22. Futures are available on three-month T-bills with a contract size of $1 million. If you take a long position at 96.22 and later sell the contracts at 96.87, what is the total net gain or loss on this transaction?

23. A bank customer will be going to London in June to purchase £100 000 in new inventory. The current spot and futures exchange rates are as follows:

Exchange Rates (Dollars/Pound)

Period	Rate
Spot	1.5342
March	1.6212
June	1.6901
September	1.7549
December	1.8416

The customer enters into a position in June futures to fully hedge their position. When June arrives, the actual exchange rate is $1.725 per pound. How much did they save?

24. Consider a put contract on a Canada bond with an exercise price of $101\frac{12}{32}$. The contract represents $100 000 of bond principal and has a premium of $750. The actual Canada bond price falls to $98\frac{16}{32}$ at the expiration. What is the gain or loss on the position?

25. A swap agreement calls for Durbin Industries to pay interest annually, based on the one-year T-bill rate, currently 6%, plus 1.5%. In return, Durbin receives interest at a rate of 6% on a fixed-rate basis. The notional principal for the swap is $50 000. What is Durbin's net interest for the year after entering into the agreement?

DATA ANALYSIS PROBLEMS

1. **Real-Time Data Analysis** Go to the St. Louis Federal Reserve FRED database and find data on the U.S. dollar/euro exchange rate (DEXUSEU). Suppose that one month ago you entered into a forward contract to sell 10 million euros at the forward rate equivalent to the spot rate one month ago. What exchange rate was specified on the forward contract? (Find the most recent data available and then use the data from one month prior.) What will be the payment in dollars when the contract is executed? In hindsight, would you have been better off or worse off by not entering into the forward contract?

2. **Real-Time Data Analysis** Go to the St. Louis Federal Reserve FRED database and find data on crude oil prices (DCOILWTICO). Suppose that one month ago you entered into a forward contract to sell 1 million barrels of oil at the forward rate equivalent to the spot rate one month ago. What price was specified on the forward contract? (Find the most recent data available and then use the data from one month prior.) What will be the payment when the contract is executed? In hindsight, would you have been better off or worse off by not entering into the forward contract?

PART 5

Central Banking and the Conduct of Monetary Policy

Crisis and Response: The Bank of Canada Response to the Global Financial Crisis and the Coronavirus Pandemic

When the Bank of Canada was confronted with what former Federal Reserve chairperson Alan Greenspan described as a "once-in-a-century credit tsunami," it resolved to come to the rescue. Starting in September 2007, the Bank of Canada enacted monetary policy at a rapid rate by reducing the overnight interest rate. Specifically, the Bank lowered the overnight rate target in December by 25 basis points (0.25 percentage points) from 4.50% to 4.25% and eventually drove the overnight rate target down to 0.25% by April 2009.

At the same time, the Bank of Canada implemented large liquidity injections into the credit markets to try to get them lending again. Over the course of the crisis, the Bank broadened its provision of liquidity to the financial system well outside of its traditional lending to depository institutions. The number of new Bank of Canada instruments over the course of the crisis spawned a whole new terminology, including term PRAs, quantitative easing, and credit easing, making the Bank of Canada sound like the military with code-named initiatives and weapons. Like the military, the Bank of Canada was fighting a war, although its weapons were financial rather than guns, tanks, or aircraft.

Once the severity of the coronavirus pandemic became clear, the Bank of Canada again went to war. In March 2020, it again lowered the overnight interest rate to zero and re-established many of the lending programs it had set up during the global financial crisis. In addition, the Bank of Canada introduced a number of lending programs to lend to new sectors of the economy, including small businesses, corporations, and provincial and municipal governments.

The recent global financial crisis and the coronavirus pandemic demonstrate the importance of central banks like the Bank of Canada to the health of the financial system and the economy. Chapter 14 outlines what central banks are trying to achieve, what motivates them, and how they are set up. Chapter 15 describes how the money supply is determined. In Chapter 16, we look at the tools that central banks like the Bank of Canada have at their disposal and how they use them. Chapter 17 extends the discussion of the conduct of monetary policy to focus on the broader picture of central banks' strategies and tactics.

14 Central Banks and the Bank of Canada

Learning Objectives

14.1 Recognize the historical context of the development of the Bank of Canada.

14.2 Describe the key features and functions of the Bank of Canada.

14.3 Assess the degree of independence of the Bank of Canada.

14.4 Summarize the arguments for and against the independence of the Bank of Canada.

14.5 Identify the ways in which the theory of bureaucratic behaviour can help explain central bank actions.

14.6 Assess the degree of independence of other major central banks around the world.

Preview

Among the most important players in financial markets throughout the world are central banks, the government authorities in charge of monetary policy. Central banks' actions affect interest rates, the amount of credit available, and the money supply, all of which have direct impacts not only on financial markets but also on aggregate output and inflation. To understand the role that central banks play in financial markets and the overall economy, we need to understand how these organizations work. Who controls central banks and determines their actions? What motivates their behaviour? Who holds the reins of power?

In this chapter, we look at the institutional structure of major central banks and focus particularly on the Bank of Canada, Canada's central bank, often just called the Bank. We start by examining the elements of the Bank of Canada's institutional structure that determine where the true power within the Bank of Canada lies. By understanding who makes the decisions, we will have a better idea of how they are made. We then examine the reasons behind central bank behaviour and whether it is a good idea to make central banks independent by insulating them from politicians. Finally, we look at the structure and independence of other major central banks, particularly the Federal Reserve System in the United States—the most important central bank in the world—and the European Central Bank. With this context in place, we will be prepared to comprehend the actual conduct of monetary policy described in the following chapters.

14.1 ORIGINS OF THE BANK OF CANADA

LO 14.1 Recognize the historical context of the development of the Bank of Canada.

The devastation of the Great Depression was of fundamental importance in the creation of the Bank of Canada. From 1929 to 1933, Canadian real GDP fell by almost 30%, and the unemployment rate increased sevenfold from less than 3% to close to 20%. The Great Depression not only involved the largest decline in the level of economic activity in the history of Canada but was also followed by an extremely slow recovery. Being such a cataclysmic event, the Great Depression contributed to significant changes in government policy, including fiscal policy, monetary policy, banking policy, and international policy.

In particular, since the depth of the Great Depression was blamed on the operation of the monetary system, in 1933 the federal Conservative government established a royal commission to study the problems of the Great Depression. Based on a

Global Establishment of Selected Central Banks

Canada did not have a central bank for almost the first 70 years after Confederation. By the time the Bank of Canada began operations in 1935, most other countries already had a central bank. In fact, as you can see in the table, Sweden and the United Kingdom created their central banks back in the seventeenth century. These early central banks, however, were initially privately owned and gradually evolved into modern publicly owned central banks.

Country	Year Central Bank Was Established
Sweden	1656
United Kingdom	1694
France	1800
Belgium	1850
Germany	1875
Japan	1882
Italy	1893
Switzerland	1905
United States	1913
Canada	1935

Source: Forrest H. Capie, Terence C. Mills, and Geoffrey E. Wood, "Central Bank Dependence and Inflation Performance: An Exploratory Data Analysis," in *Varieties of Monetary Reform: Lessons and Experiences on the Road to Monetary Union*, ed. Pierre L. Siklos (Dordrecht, The Netherlands: Kluwer Academic Publishers, 1994), 95–132.

recommendation of the royal commission, Parliament passed the Bank of Canada Act in 1934, and the newly founded Bank of Canada started operations on March 11, 1935. By this time, most other countries already had a central bank. (See the Global box "Establishment of Selected Central Banks.") Although the primary motivation for the formation of the Bank of Canada was economic (or monetary), other motives within the government included the need for Canada to reflect its growing political independence from Britain and the need to coordinate its international economic policy.

Initially the Bank of Canada was a private institution but was nationalized in 1938, and it is now a national institution with headquarters in Ottawa. The Bank also has regional offices in Toronto (for Ontario), Vancouver (for British Columbia and the Yukon), Calgary (for the Prairies, Northwest Territories, and Nunavut), Montréal (for Quebec), and Halifax (for Atlantic Canada). In 2002, the Bank also established a post in New York to enhance communication with the world's largest financial community. Unlike a private bank that operates in pursuit of profit, the Bank of Canada is responsible for the country's monetary policy and for the regulation of Canada's deposit-based financial institutions.

14.2 FORMAL STRUCTURE OF THE BANK OF CANADA

LO 14.2 Describe the key features and functions of the Bank of Canada.

The overall responsibility for the operation of the Bank of Canada rests with a **Board of Directors**, which consists of 15 members—the governor, the senior deputy governor, the deputy minister of finance, and 12 outside directors. The Board appoints

the governor and senior deputy governor, with the government's approval, for a renewable term of seven years. The outside directors are appointed by the minister of finance, with Cabinet approval, for a three-year term, and they are required to come from all regions of Canada, representing a variety of occupations with the exception of banking. The governor of the Bank is the chief executive officer and chair of the Board of Directors. Currently, the governor of the Bank of Canada is Tiff Macklem. (See the FYI box "The Political Environment and the Bank of Canada.") Tiff Macklem replaced Stephen Poloz (the former Bank of Canada governor) on June 3, 2020.

In 1994, the Board of Directors made some changes in the internal organization of the Bank. Most prominently, the Board established a new senior decision-making authority within the Bank called the **Governing Council**. The Council is chaired by the governor and is composed of the senior deputy governor and four deputy governors. Since this change, the six members of the Governing Council of the Bank collectively assume responsibility for the Bank's quarterly *Monetary Policy Report*, published in January, April, July, and October. This system of "collective responsibility" ensures that the Bank's governor is not personally identified with the Bank's policy.

The Functions of the Bank of Canada

In the words of the preamble of the Bank of Canada Act, the functions of the Bank of Canada are

> to regulate credit and currency in the best interests of the economic life of the nation, to control and protect the external value of the national monetary unit and to mitigate by its influence fluctuations in the general level of production, trade, prices and employment, so far as may be possible within the scope of monetary action, and generally to promote the economic and financial welfare of Canada.

FYI The Political Environment and the Bank of Canada

Since the inception of the Bank of Canada, there have been eight governors:

1935–1954, Graham Towers
1955–1961, James Coyne
1961–1973, Louis Rasminsky
1973–1987, Gerald Bouey
1987–1994, John Crow
1994–2000, Gordon Thiessen
2001–2007, David Dodge
2008–2013, Mark Carney
2013–2020, Stephen Poloz
2020– Tiff Macklem

It is interesting to note that during the same period, Canadians went to the polls more than 20 times to elect a federal government. The Bank of Canada is not completely independent from the government. For example, the government can directly influence the Bank by not renewing the governor's appointment when it expires, as the Liberal government did in 1994 when it didn't renew the appointment of John Crow, who was appointed by the Conservative government in 1987.

This is a vague mandate, leaving a lot of room for interpretation. To explore this subject, we discuss four main areas of responsibility of the Bank of Canada as they are mentioned in the Bank's webpage:

- Currency
- Funds management
- Financial system
- Monetary policy

Currency Before the creation of the Bank of Canada, the federal government and the early banks issued notes designed to circulate as currency. The day it began operations, the Bank replaced the outstanding issue of federal government notes, and provision was also made for the gradual removal of notes issued by banks. By 1945, the Bank had a monopoly over note issue. Although the original Bank Act required the Bank to redeem its notes in gold, this provision was never used. In fact, it was removed with the 1967 revision of the Bank Act, thereby providing the Bank with unlimited powers to issue legal tender.

The Bank also conducts ongoing research, working closely with private-sector partnerships and note-issuing authorities in other countries, to improve cost-effectiveness, increase the durability of bank notes, and reduce counterfeiting. In its role as provider of paper money, the Bank's overall objective is to preserve the integrity and safety of Canadian currency in the most economical and efficient manner possible.

Funds Management In its role as the federal government's "fiscal agent," the Bank of Canada provides debt-management services for the federal government, such as advising on borrowings, managing new debt offerings, and servicing outstanding debt. Before 1995, these services were provided for all the federal government's debt. In 1995, however, a special agency of the Department of Finance was created, known as Canada Investment and Savings, to be responsible for the federal government's debt held by individuals, commonly known as retail debt.

Canada Investment and Savings handles Government of Canada securities such as Treasury bills and marketable bonds and is also responsible for the development of new investment products and marketing initiatives. The Bank of Canada, however, continues to be responsible for all the government's securities after they are issued, administering millions of bondholder accounts and making payments on behalf of the federal government for interest and debt redemption.

In its role as fiscal agent, the Bank of Canada also manages the government's foreign exchange reserves held by the **Exchange Fund Account** of the Department of Finance. In particular, the Bank assists the Department of Finance in investing these foreign reserves and in borrowing when necessary to maintain an adequate level of reserves. The Bank also engages in international financial transactions, on behalf of the government, in order to influence exchange rates. (We discuss the Bank's foreign exchange interventions more formally in Chapter 19.)

Financial System The Bank of Canada serves as the lender of last resort if a deposit-taking financial institution faces a liquidity crisis. Because of its unique power to create base money, the Bank can ease the liquidity problems of any financial institution by extending advances and therefore deter bank runs and panics. **Base money**

(also called **monetary base**) consists of the monetary liabilities of the central bank and, as you will see in the next chapter, is an important part of the money supply, because changes in it lead to multiple changes in the money supply. Of course, lender-of-last-resort lending is closely coordinated with the two federal regulatory agencies that are set up specifically to regulate financial institutions—the Office of the Superintendent of Financial Institutions Canada and the Canada Deposit Insurance Corporation.[1] Moreover, such lending is done judiciously, explicitly considering the effects on other financial institutions, the money supply, and government policy.

The Bank of Canada also plays a central role in Canada's national payments system. This is essentially an electronic system that clears and settles payments and transactions including securities and foreign exchange, currently handling close to 25 times our gross domestic product per year. Although this system is operated by Payments Canada (formerly, the Canadian Payments Association), federal legislation that came into force in 1996 gave the Bank explicit responsibility for the regulatory oversight of this system. The Bank's main concern is whether problems that affect one participant in the clearing and settlement system will spread to other participants.

Finally, the Bank of Canada acts as the holder of deposit accounts of the federal government, the directly clearing members of Payments Canada, international organizations such as the International Monetary Fund, and other central banks. As the federal government's banker, the Bank is also responsible for the government's operating accounts. In this role, as you will see in Chapter 16, the Bank shifts government balances between the government's transactions account with the Bank of Canada and the government's non-transactions accounts with the deposit-taking institutions, using twice-daily auctions of government term deposits.

Monetary Policy The Bank of Canada employs such tools as **open market buyback operations** (the purchase and sale of government securities that affect both interest rates and the amount of reserves in the banking system) and, to a lesser extent, the shifting of government balances between it and the directly clearing members of Payments Canada to implement changes in the money supply. The Bank's ultimate objective is to keep inflation low. The Bank has a staff of professional economists, which provides economic analysis that the Board of Directors uses in making its decisions. (See the FYI box "Role of the Bank's Research Staff.")

The Bank's goal of low inflation is closely related to the goal of steady economic growth, because businesses are more likely to invest in capital equipment to increase productivity and economic growth when inflation is low. Low inflation is also desirable because it protects the purchasing power of pensioners and those on fixed incomes.

Although the Bank determines monetary policy, in the following section you will learn that the ultimate responsibility for policy rests with the government, since it is the government that must answer to Parliament. This system of "joint responsibility" dates back to 1967, when the Bank of Canada Act was amended to give responsibility for monetary policy to the government.

[1]The Office of the Superintendent of Financial Institutions was created in 1987 to succeed the Department of Insurance and the Inspector General of Banks, whereas the Canada Deposit Insurance Corporation was created by an act of Parliament in 1967 to insure deposits, currently up to $100 000 per account, of member deposit-taking institutions.

FYI Role of the Bank's Research Staff

The Bank of Canada is the largest employer of economists in Canada. What do all these economists do?

The most important task of the Bank's economists is to follow the incoming data from government agencies and private-sector organizations on the economy and provide guidance to policymakers on where the economy may be heading and what the impact of monetary policy actions on the economy might be. Moreover, the Bank's economists maintain large econometric models (models whose equations are estimated with statistical procedures) that help them produce forecasts of the national economy, and they brief the governor and the senior management of the Bank on their forecasts for the Canadian economy.

Because of the increased influence of developments in foreign countries on the Canadian economy, the research staff produces reports on major foreign economies. They also conduct research on developments in the foreign exchange market because of its growing importance in the monetary policy process and to support the activities of the Bank's foreign exchange desk.

Staff economists also engage in basic research on the effects of monetary policy on output and inflation, developments in the labour markets, international trade, international capital markets, banking and other financial institutions, financial markets, and regional economies, among other topics. This research is published widely in academic journals and in Bank of Canada publications. Bank of Canada publications, such as the *Annual Report* (published in March each year), the *Quarterly Financial Reports*, the *Bank of Canada Review* (published in May and November), the *Financial System Review* (published in June and December), the *Bank of Canada Banking and Financial Statistics* (published monthly), and the *Monetary Policy Report* (published quarterly) are good sources of supplemental material for money and banking students.

Another important activity of the research staff is in the public education area. Staff economists are called on frequently to make presentations to the public.

14.3 HOW INDEPENDENT IS THE BANK OF CANADA?

LO 14.3 Assess the degree of independence of the Bank of Canada.

When we look, in the next three chapters, at how the Bank of Canada conducts monetary policy, we will want to know why the Bank decides to take certain policy actions but not others. To understand its actions, we must understand the incentives that motivate the Bank's behaviour. How free is the Bank of Canada from the whims of the government? Do economic, bureaucratic, or political considerations guide it? Is the Bank of Canada truly independent of outside pressures?

Stanley Fischer, who was a professor at MIT and much later a vice-chair of the U.S. Federal Reserve, has defined two different types of independence of central banks: **instrument independence**, the ability of the central bank to set monetary policy instruments, and **goal independence**, the ability of the central bank to set the goals of monetary policy. Unlike the U.S. Federal Reserve and the European Central Bank, which have both types of independence and are remarkably free of political pressures that influence other government agencies, the Bank of Canada has instrument independence but not goal independence.

The Bank of Canada's degree of independence has evolved over time, in part because of changing circumstances and in part because a clear division of authority was not established in the original Bank of Canada Act. Initially, the Bank of Canada was privately owned, with about 12 000 individual shareholders, and so was largely free of political pressures. It was also free of private interference because of regulations

regarding who could hold how much stock in the Bank. The Conservative government in office at the time of the Bank's creation believed that the Bank should possess a large share of the responsibility for the development of monetary policy.

Significant changes in the balance of power occurred with the election of a majority Liberal government in the fall of 1935. The Liberals moved the Bank of Canada in the direction of public ownership, culminating in its complete nationalization by 1938. This was done to further isolate the Bank from the pressures of the private system. However, the nationalization of the Bank of Canada did tilt the balance of authority for monetary policy back toward the federal Cabinet and Parliament. The Liberals believed that the Bank of Canada should have discretion in internal management and in implementing monetary policy, yet the policies being implemented should be in harmony with the views of the government.

In June 1954, the first governor of the Bank of Canada, Graham Towers, retired after 19 years of service. Towers was replaced by James Coyne. (See the FYI box "The Political Environment and the Bank of Canada.") Like many Canadians at the time, Coyne was concerned about the level of foreign ownership and the increase in consumer prices during the Korean War. He was convinced that the solution to these problems was a tighter monetary policy stance. A decline in the supply of bank reserves would increase interest rates, thereby raising national savings and reducing Canada's dependence on foreign capital inflows. This tighter monetary policy led to rising unemployment and weak output growth. As a result, the Bank of Canada's policies were criticized in the popular press and by most of the academic community. In fact, 30 economists signed a letter to the minister of finance calling for Coyne's dismissal.[2]

James Coyne, however, did not change his mind. He was convinced that there was no long-term relationship between inflation and unemployment, despite the publication of a famous paper in 1958 by the British economist A. W. Phillips showing that higher inflation was typically associated with a lower unemployment rate.[3] Although the Phillips hypothesis attracted widespread support in Canada and was replicated with Canadian data, Coyne was anticipating the pathbreaking work by Milton Friedman and Edmund Phelps on the vertical Phillips curve, to be published 10 years later in 1968.[4] But as a former governor of the Bank of Canada, Gordon Thiessen, put it in a recent speech,

> There was one critical area, however, where Coyne and many other policy analysts, both within and outside the Bank, appear to have been misguided. Those who had questioned the effectiveness of monetary policy in earlier years had failed to appreciate that it was likely to be much stronger than fiscal policy under a flexible exchange rate system, especially when capital was highly mobile. The large capital movements triggered by any change in interest rates would put significant pressure on the exchange rate, amplifying the effects of monetary policy while undercutting the effects of any opposing fiscal policy. Coyne did not realize that, for similar reasons, it was unlikely that a tighter monetary policy would ever raise national savings or reduce foreign investment inflows.[5]

[2]See H. S. Gordon, *The Economists Versus the Bank of Canada* (Toronto: Ryerson Press, 1961), v–vi.

[3]A. W. Phillips, "The Relation Between Unemployment and the Rate of Change of Money Wages in the United Kingdom," *Economica* 24 (1958): 283–299.

[4]Milton Friedman, "The Role of Monetary Policy," *American Economic Review* 58 (1968): 1–17; Edmund Phelps, "Money-Wage Dynamics and Labour Market Equilibrium," *Journal of Political Economy* 76 (1968): 678–711.

[5]Gordon Thiessen, "Can a Bank Change? The Evolution of Monetary Policy at the Bank of Canada 1935–2000," Lecture to the Faculty of Social Science, University of Western Ontario (October 17, 2000). Available online courtesy of the Bank of Canada.

To avoid accepting blame for this state of affairs, the Liberal government in office reversed its previous stand and disavowed its position that the government was responsible for monetary policy. The government was criticized for this reversal and the new hands-off approach to monetary policy. When the government changed hands on June 10, 1957, and Donald Fleming became the federal minister of finance, the stage was set for a confrontation with the governor of the Bank of Canada, James Coyne. After a long and acrimonious debate, the so-called Coyne Affair was resolved in 1961 with the resignation of Governor Coyne.

On July 24, 1961, Louis Rasminsky accepted the position as the third governor of the Bank of Canada on the condition that the relationship between the government and the Bank of Canada be clearly defined. Soon after he assumed office, he issued a public statement containing two main principles reflecting his views on that relationship:

> (1) In the ordinary course of events, the Bank has the responsibility for monetary policy, and (2) if the government disapproves of the monetary policy being carried out by the Bank, it has the right and the responsibility to direct the Bank as to the policy which the Bank is to carry out.[6]

A royal commission, whose appointment was partly initiated by the Coyne Affair, accepted Governor Rasminsky's views regarding the relationship between the Bank of Canada and the government. The commission recommended a system of joint responsibility under which the Bank of Canada has considerable autonomy in the conduct of day-to-day monetary policy but the government must accept full responsibility for the policy being followed.

The Bank of Canada Act was amended in 1967 to confirm the joint responsibility system, and this state of affairs regarding monetary policy has generally remained in order to this day. Under this joint responsibility system, the governor of the Bank of Canada and the minister of finance, acting on behalf of the government, consult regularly, and, in the event of a serious disagreement over the conduct of monetary policy, the government has the right to override the Bank's decisions. In particular, the minister of finance can issue a directive to the Bank indicating the specific policy changes that the Bank must follow. The directive, however, must be published, indicating not only the new policy that the Bank is supposed to undertake but also the period during which it is to apply.

Hence, ultimate responsibility for monetary policy rests with the democratically elected government. However, because of the consequences of issuing a directive, it is unlikely that such a directive would be issued, and none has been issued to date.

The Changing Face of the Bank of Canada

The legislation governing the Bank of Canada's responsibility for monetary policy has remained relatively unchanged since the Bank Act of 1967. Over the past decade, however, the Bank has undertaken a broad range of initiatives and made significant institutional changes to the way it operates. The impetus for change came from the interaction of experience and economic theory, the desire to explain and build confidence in the Bank's actions, and (to a lesser extent) from technological change and globalization.

[6]Bank of Canada, *Annual Report*, 1961, p. 3.

Clarifying objectives is an important starting point for any successful monetary policy framework. The legislation by which the Bank of Canada is governed does not facilitate a clear understanding of objectives. For example, the preamble to legislation governing the Bank refers to multiple, and potentially inconsistent, policy objectives. Over the past few decades, however, the performance of the Canadian economy, together with the evolution of economic theory, led to the view that price stability is the most important goal of monetary policy. As John Crow put it, in his January 1988 Hanson Memorial Lecture at the University of Alberta,

> Theory and experience—much of this experience not overly cheerful but certainly instructive—both point to a very clear answer. Monetary policy should be conducted so as to achieve a pace of monetary expansion that promotes stability in the value of money. This means pursuing a policy aimed at achieving and maintaining stable prices.[7]

The Hanson lecture was designed to explicitly identify price stability as the objective of Bank of Canada policy. This message was reinforced on February 26, 1991, through a joint announcement by the governor of the Bank and the minister of finance regarding the establishment of formal inflation targets (to be discussed in detail in Chapter 17). Clearly defined targets can be appealing as an institutional way of improving the macroeconomic outcomes of monetary policy.

The effectiveness of such an institutional framework, however, depends on two fundamental requirements: "independence" and "accountability." Although the Bank of Canada has not been given "goal independence"—the goal of price stability has been set jointly by the Bank and the Ministry of Finance—over the past decade the Bank has placed increased emphasis on its responsibility to achieve the goal of price stability and its greater freedom to take whatever action is needed to do so. This freedom is referred to as "operational" (or "instrument") independence, and, although it hasn't been explicitly legislated, it exists in practice. Increased operational independence has also raised the standards for accountability. As already noted, for example, the governor of the Bank has delegated the authority for monetary policy decisions to a committee—the six-member internal Governing Council. Moreover, it has become a de facto standard for the governor to appear before a parliamentary committee following the release of the Bank's *Monetary Policy Report*.

The Bank of Canada has also improved its communications activities by moving toward greater transparency in its operations and objectives. In particular, the Bank has noticeably increased the number of its press conferences, press releases, and speeches, and also reorganized its regional offices in 1996–1997, with the objective of improving communication, transparency, and its assessment of economic conditions across Canada. For example, the Bank's regional offices, by maintaining contact with provincial governments, industries, and the general public, present quarterly "grassroots" assessments of current and prospective economic developments to the Bank's Governing Council—information that complements economic projections prepared by the Bank's staff. Finally, the Bank of Canada maintains a comprehensive website (at www.bankofcanada.ca) to disseminate information regarding financial statistics, publications, the transmission of monetary policy, and Bank-related material.[8]

[7]John Crow, "The Work of Canadian Monetary Policy," The Eric J. Hanson Memorial Lecture, University of Alberta, Edmonton, Alberta. *Bank of Canada Review* (1988): 3–17.

[8]An inside view of how the Bank interacts with the public and the politicians can be found in John Crow, *Making Money: An Insider's Perspective on Finance, Politics, and Canada's Central Bank* (Toronto: John Wiley, 2002).

The direction taken in the recent evolution of the monetary policy framework in Canada has been heavily influenced by the role that the institutional monetary structure plays in influencing the monetary conduct.[9] As the Bank's former governor Gordon Thiessen put it in his October 17, 2000, speech to the Faculty of Social Science of the University of Western Ontario,

> The Bank tries to work with markets, rather than against them, to avoid surprising them with unexpected actions. Greater transparency facilitates the policy-transmission process by conditioning market expectations, and helps avoid unnecessary confusion about the reasons for our actions.[10]

From Opaqueness to Accountability and Transparency The Bank of Canada did not release information regarding its guidelines or resulting consequences during the 1960s and 1970s. As a general rule, the Bank wanted to remain nontransparent, keeping its monetary policy approaches secret. The use of multiple instruments and goals, vague statements about its policies, and moral suasion with banks resulted in an unnecessary level of secrecy. Due to this instrument and goal opaqueness, the public was not able to question or comprehend the Bank's actions, allowing the accountability of the Bank of Canada to deteriorate.

Over the past decade, however, the Bank of Canada has rejected multiple policy instruments by adopting the overnight interest rate as the centrepiece of its monetary policy implementation and by focusing on an explicit inflation-control target. The goal of the Bank's current monetary policy is to keep the inflation rate within a target range of 1% to 3%, with the midpoint of the inflation target range, 2%, being the most desirable outcome; the 1% to 3% target range for inflation applies until the end of 2026. By announcing a target overnight interest rate, establishing planned dates for policy changes, and implementing the Governing Council for making decisions and disclosing information, the Bank of Canada has moved toward greater accountability and transparency.

The Bank's move toward accountability and transparency was motivated by a number of recent trends in society and the economy. For example, the shift from fixed to flexible exchange rates removed an obstacle to openness. The experience with inflation in the 1970s and the recession in the early 1980s allowed the Bank of Canada to reconsider its analysis of inflation and its costs. In addition, economic comprehension about interest rates, developments about targets and inflation, and greater openness in the government also contributed to the accountability and transparency of the Bank of Canada.[11]

[9]For a discussion of similar changes implemented by other central banks around the world, see Graydon Paulin, "The Changing Face of Central Banking in the 1990s," *Bank of Canada Review* (Summer 2000): 3–13.

[10]Gordon Thiessen, "Can a Bank Change? The Evolution of Monetary Policy at the Bank of Canada 1935–2000." Lecture to the Faculty of Social Science, University of Western Ontario (October 17, 2000).

[11]For more details regarding the Bank's recent move toward openness and accountability, see John Chant, "The Bank of Canada: Moving Towards Transparency," *Bank of Canada Review* (Spring 2003): 5–13. Also, regarding Canada's experience with central banking, see Michael D. Bordo and Angela Redish, "70 Years of Central Banking: The Bank of Canada in an International Context, 1935–2005," *Bank of Canada Review* (Winter 2005–2006): 7–14, and David Laidler, "Free Banking and the Bank of Canada," *Bank of Canada Review* (Winter 2005–2006): 15–24.

14.4 SHOULD THE BANK OF CANADA BE INDEPENDENT?

LO 14.4 Summarize the arguments for and against the independence of the Bank of Canada.

As we have seen, the Bank of Canada is probably the most independent government agency in Canada. Every few years, the question arises whether the independence given to the Bank of Canada should be curtailed. Politicians who strongly oppose a Bank policy often want to bring it under their supervision in order to impose a policy more to their liking. Should the Bank of Canada be independent, or would we be better off with a central bank under the control of the government?

The Case for Independence

The strongest argument for an independent Bank of Canada rests on the view that subjecting the Bank to more political pressures would impart an inflationary bias to monetary policy. In the view of many observers, politicians in a democratic society are short-sighted because they are driven by the need to win their next election. With this as the primary goal, they are unlikely to focus on long-run objectives, such as promoting a stable price level. Instead, they will seek short-run solutions to problems like high unemployment and high interest rates, even if the short-run solutions have undesirable long-run consequences. For example, we saw in Chapter 5 that high money growth might lead initially to a drop in interest rates but might also cause an increase later, as inflation heats up. Would a Bank of Canada under the control of the government be more likely to pursue a policy of excessive money growth when interest rates are high, even though such a policy would eventually lead to inflation and even higher interest rates in the future? The advocates of an independent central bank say yes. They believe that a politically insulated central bank is more likely to be concerned with long-run objectives and thus more likely to defend a sound dollar and a stable price level.

A variation on the preceding argument is that the political process in Canada could lead to a **political business cycle**, in which just before an election, expansionary policies are pursued to lower unemployment and interest rates. After the election, the bad effects of these policies—high inflation and high interest rates—come home to roost, requiring contractionary policies that politicians hope the public will forget before the next election. Although the issue has not been completely settled, research indicates that there is no credible evidence that such a political business cycle exists in Canada. (See the FYI box "Economics and Politics.")

Putting the Bank of Canada under the control of the government is also considered dangerous because the Bank can be used to facilitate government financing of large budget deficits by its purchases of government bonds. Such open market purchases by the Bank of Canada increase the money supply and lead to inflation. Government pressure on the Bank of Canada to "help out" might lead to a more inflationary bias in the economy. An independent Bank of Canada is better able to resist this pressure from the government.

Another argument for central bank independence is that control of monetary policy is too important to leave to politicians, a group that has repeatedly demonstrated a lack of expertise at making hard decisions on issues of great economic importance, such as reducing the budget deficit or reforming the banking system. This argument can also be stated in terms of the principal–agent problem discussed in Chapters 8, 9, and 10. Both the Bank of Canada and politicians are agents of the public (the principals), and both politicians and the Bank of Canada have incentives to act in their own interest

FYI Economics and Politics

Recent politico-institutional approaches to the theory of economic policy emphasize the incentives of rational and maximizing policymakers in explaining movements in macroeconomic variables. Central to this perspective is the general assumption that policymakers respond to incentives and constraints just like all other economic agents. As a consequence, the actual policies of government give rise to political cycles, which on the basis of the primary motivational force involved are distinguished into opportunistic and partisan cycles.

In opportunistic (or electoral) business cycle models, politicians maximize their popularity or their probability of re-election by following pre-election expansionary fiscal policies in order to please the fiscally knowledgeable voters. As a consequence, their actual policies give rise to **electoral business cycles**—that is, persistent cyclical patterns of key policy and target variables across electoral terms, regardless of the political orientation of the incumbent government. In particular, models of opportunistic cycles predict pre-election high growth and low unemployment, increasing inflation around the election time, and a post-election contraction, regardless of the political party in power.

In the more accepted partisan cycle models, politicians are ideological; that is, they represent the interests of different pressure groups and, when in office, follow policies that are favourable to their supporting groups. For example, the left-wing parties pursue expansionary policies in order to reduce unemployment, while the right-wing parties tend to induce post-election contractions in economic performance in order to reduce inflation. The outcome, therefore, is **partisan business cycles**—that is, systematic and permanent differences in macroeconomic outcomes that differ by political party.

Work by Apostolos Serletis and Panos Afxentiou of the University of Calgary indicates that there is no credible evidence that such a political business cycle exists in Canada.[12]

[12]Apostolos Serletis and Panos C. Afxentiou, "Electoral and Partisan Cycle Regularities in Canada," *Canadian Journal of Economics* 31 (1998): 28–46.

rather than in the interest of the public. The argument supporting Bank of Canada independence is that the principal–agent problem is worse for politicians than for the Bank of Canada because politicians have fewer incentives to act in the public interest.

Indeed, some politicians may prefer an independent Bank of Canada because it can be used as a public "whipping boy" to take some of the heat off their backs. It is possible that a politician who in private opposes an inflationary monetary policy will be forced to support such a policy in public for fear of not being re-elected. An independent Bank of Canada can pursue policies that are politically unpopular yet in the public interest.

The Case Against Independence

Proponents of a Bank of Canada under the control of the government argue that it is undemocratic to have monetary policy (which affects almost everyone in the economy) controlled by an elite group responsible to no one. The current lack of accountability of the Bank of Canada has serious consequences: If the Bank performs badly, no provision is in place for replacing members (as there is with politicians). True, the Bank of Canada needs to pursue long-run objectives, but elected government officials vote on long-run issues also (foreign policy, for example). If we push the argument further that policy is always performed better by elite groups like the Bank of Canada, we end up with such conclusions as that the Canada Revenue Agency should set tax policies with no oversight from the government. Would you advocate this degree of independence for the Canada Revenue Agency?

The public holds the government responsible for the economic well-being of the country, yet it lacks control over the government agency that may well be the most important actor in determining the health of the economy. In addition, to achieve a cohesive program that will promote economic stability, monetary policy must be coordinated with fiscal policy (the management of government spending and taxation). Only by placing monetary policy under the control of the politicians who also control fiscal policy can these two policies be prevented from working at cross-purposes.

No consensus has yet been reached on whether Bank of Canada independence is a good thing, although public support for independence of the central bank seems to be growing in both Canada and abroad. As you might expect, people who like the Bank's policies are more likely to support its independence, whereas those who dislike its policies advocate a less independent Bank of Canada.

Central Bank Independence and Macroeconomic Performance Throughout the World

We have seen that advocates of an independent central bank believe that macroeconomic performance will be improved by making the central bank more independent. Empirical evidence seems to support this conjecture: When central banks are ranked from least independent to most independent, inflation performance is found to be the best for countries with the most independent central banks. Although a more independent central bank appears to lead to a lower inflation rate, this is not achieved at the expense of poorer real economic performance. Countries with independent central banks are no more likely to have high unemployment or greater output fluctuations than are those countries with less independent central banks.

14.5 EXPLAINING CENTRAL BANK BEHAVIOUR

LO 14.5 Identify the ways in which the theory of bureaucratic behaviour can help explain central bank actions.

One view of government bureaucratic behaviour is that bureaucracies serve the public interest (this is called the *public interest view*). Yet some economists have developed a theory of bureaucratic behaviour that emphasizes other factors that influence the operation of bureaucracies. The *theory of bureaucratic behaviour* suggests that the objective of a bureaucracy is to maximize its own welfare, just as a consumer's behaviour is motivated by the desire to maximize personal welfare and a firm's behaviour is motivated by the desire to maximize profits. The welfare of a bureaucracy is related to its power and prestige. Thus this theory suggests that an important factor affecting a central bank's behaviour is its wish to increase its power and prestige.

What predictions does this view of a central bank like the Bank of Canada suggest? One is that the Bank of Canada will fight vigorously to preserve its autonomy, a prediction verified time and time again as the Bank of Canada has continually counterattacked attempts to control its functions. Another prediction is that the Bank of Canada will try to avoid conflict with powerful groups that might threaten to curtail its power and reduce its autonomy. The Bank's behaviour may take several forms. One possible factor explaining why the Bank is sometimes slow to increase interest rates and thus smooth out their fluctuations is that it wishes to avoid a conflict with the government over increases in interest rates.

The theory of bureaucratic behaviour seems applicable to the Bank of Canada's actions, but we must recognize that this view of the Bank of Canada as being solely concerned with

its own self-interest is too extreme. Maximizing one's welfare does not rule out altruism. (You might give generously to a charity because it makes you feel good about yourself, but in the process you are helping a worthy cause.) The Bank of Canada is surely concerned with conducting monetary policy that is in the public interest. However, much uncertainty and disagreement often exist over what that monetary policy should be.[13] When it is unclear what policy will best serve the public interest, other motives may influence the Bank's behaviour. In these situations, the theory of bureaucratic behaviour may be a useful guide to predicting what motivates the Bank of Canada and other central banks.

14.6 STRUCTURE AND INDEPENDENCE OF FOREIGN CENTRAL BANKS

LO 14.6 Assess the degree of independence of other major central banks around the world.

In contrast to the Bank of Canada, which is a centralized unit owned by the government, central banks in other industrialized countries have a more decentralized structure. Here we examine the structure and degree of independence of four of the most important foreign central banks: the Federal Reserve System of the United States, the European Central Bank, the Bank of England, and the Bank of Japan.

Federal Reserve System

Of all the central banks in the world, the **Federal Reserve System** of the United States (also called simply **the Fed**) probably has the most unusual structure. It includes the following entities: the **Board of Governors of the Federal Reserve System**, the **Federal Reserve Banks**, the **Federal Open Market Committee (FOMC)**, the Federal Advisory Council, and around 3000 member commercial banks. Figure 14-1 outlines the relationships of these entities to one another and to the four policy tools of the Fed—open market operations, the discount rate (the U.S. equivalent of the bank rate in Canada), reserve requirements, and interest paid on excess reserves.

Board of Governors of the Federal Reserve System At the head of the Federal Reserve System is the seven-member Board of Governors, headquartered in Washington, D.C. Each governor is appointed by the president of the United States and confirmed by the Senate. To limit the president's control over the Fed and insulate the Fed from other political pressures, the governors can each serve one nonrenewable 14-year term plus part of another term, with one governor's term expiring every other January. The chair of the Board of Governors is chosen from among the seven governors and serves a four-year, renewable term. It is expected that once a new chair is chosen, the old chair resigns from the Board of Governors, even if there are many years left to their term as a governor.

Federal Reserve Banks Each of the 12 Federal Reserve districts has one main Federal Reserve bank, which may have branches in other cities in the district. The

[13]One example of the uncertainty over how best to conduct monetary policy was discussed in Chapter 3: Economists are not sure how to measure money. So even if economists agreed that controlling the quantity of money is the appropriate way to conduct monetary policy (a controversial position, as we will see in later chapters), the Bank of Canada cannot be sure which monetary aggregate it should control.

Mini-lecture

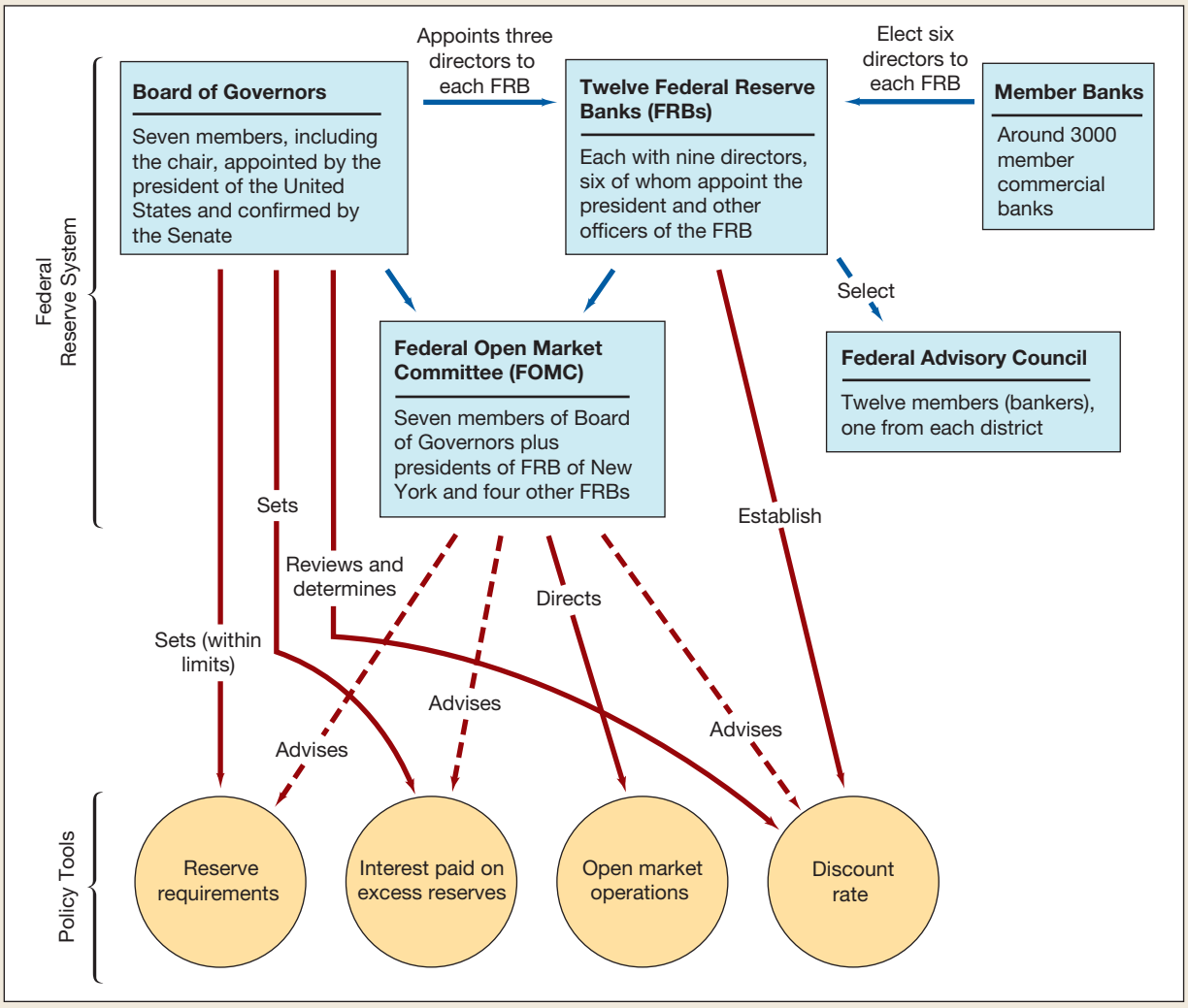

Federal Reserve System
Board of Governors
Seven members, including the chair, appointed by the president of the United States and confirmed by the Senate
Appoints three directors to each FRB
Twelve Federal Reserve Banks (FRBs)
Each with nine directors, six of whom appoint the president and other officers of the FRB
Elect six directors to each FRB
Member Banks
Around 3000 member commercial banks
Select
Federal Advisory Council
Twelve members (bankers), one from each district
Federal Open Market Committee (FOMC)
Seven members of Board of Governors plus presidents of FRB of New York and four other FRBs
Sets
Reviews and determines
Directs
Establish
Sets (within limits)
Advises
Advises
Advises
Policy Tools
Reserve requirements
Interest paid on excess reserves
Open market operations
Discount rate

FIGURE 14-1 Structure and Responsibility for Policy Tools in the Federal Reserve System

The relationships of the Federal Reserve banks, the Board of Governors of the Federal Reserve System, and the FOMC to the four policy tools of the Fed (open market operations, the discount rate, reserve requirements, and interest rate paid on excess reserves). Dashed lines indicate that the FOMC "advises" on the setting of reserve requirements and the discount rate.

locations of these districts, the Federal Reserve banks, and the Federal Reserve bank branches are shown in Figure 14-2. The three largest Federal Reserve banks in terms of assets are those of New York, Chicago, and San Francisco—combined they hold more than 50% of the assets (discount loans, securities, and other holdings) of the Federal Reserve System. The New York bank, with around one-quarter of the assets, is the most important of the Federal Reserve banks. (See the FYI box "The Special Role of the Federal Reserve Bank of New York.")

Mini-lecture

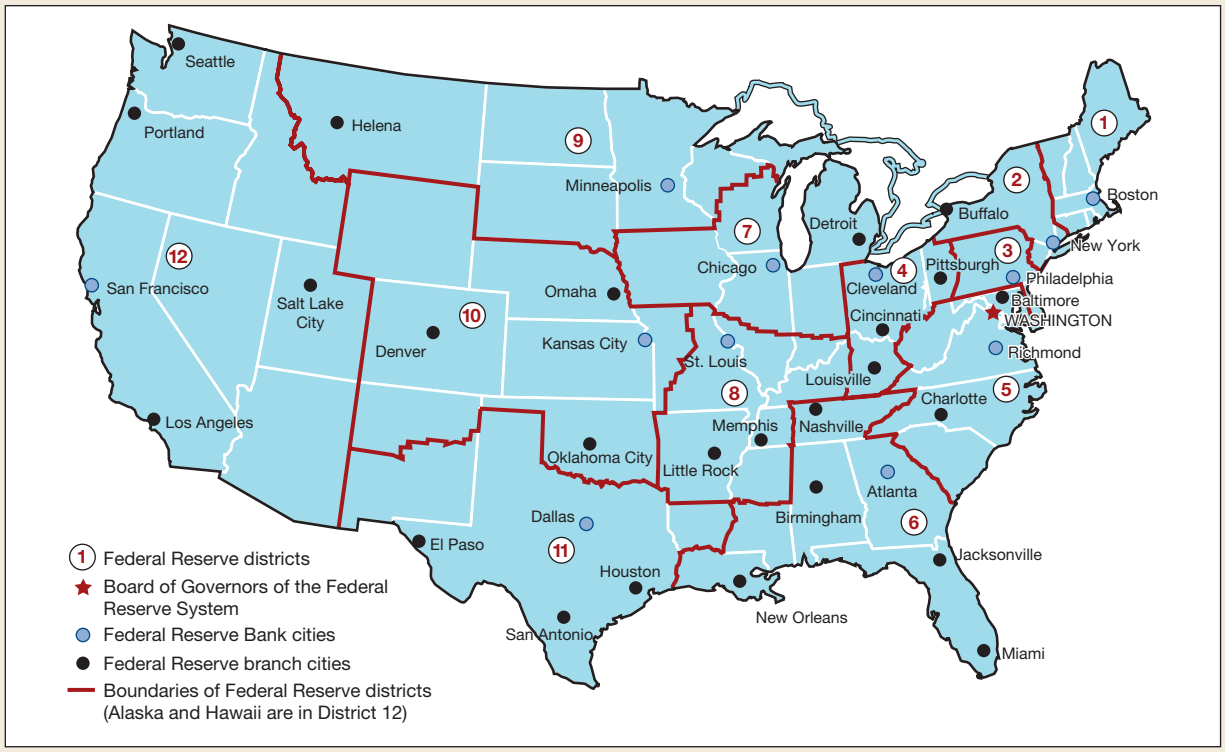


FIGURE 14-2 Federal Reserve System

The locations of the Federal Reserve districts, the Federal Reserve banks, and their branches.

Source: Federal Reserve *Bulletin*.

Federal Open Market Committee

The 12 Federal Reserve banks and the Board of Governors are actively involved in decisions concerning the conduct of monetary policy. In particular, all seven governors are voting members of the FOMC, together with the president of the Federal Reserve Bank of New York and presidents of four other Federal Reserve banks. The FOMC usually meets eight times a year (about every six weeks) and makes decisions regarding the conduct of open market operations and the setting of the policy interest rate, the **federal funds rate**, which is the interest rate on overnight loans from one bank to another. Indeed, the FOMC is often referred to as "the Fed" in the press: For example, when the media report that the Fed is meeting, they actually mean that the FOMC is meeting.

How Independent Is the Fed?

The Federal Reserve appears to be remarkably free of the political pressures that influence other government agencies in the United States. Yet it is still subject to the influence of the president of the United States, since the president appoints members to the Board of Governors. The power of the U.S. president in appointing members to the Board of Governors is limited, however. Because the term of the chairperson of the Board of Governors is not necessarily concurrent with that of the president of the United States, a president may have to deal with a

FYI The Special Role of the Federal Reserve Bank of New York

The Federal Reserve Bank of New York plays a special role in the Federal Reserve System for several reasons. First, its district contains many of the largest commercial banks in the United States, the safety and soundness of which are paramount to the health of the U.S. financial system. The Federal Reserve Bank of New York conducts examinations of bank holding companies and state-chartered member banks in its district, making it the supervisor of some of the most important financial institutions in the U.S. financial system. Not surprisingly, given this responsibility, the bank supervision group is one of the largest units of the New York Fed and is by far the largest bank supervision group in the Federal Reserve System.

The second reason for the New York Fed's special role is its active involvement in the bond and foreign exchange markets. The New York Fed houses the open market desk, which conducts open market operations—the purchase and sale of bonds—that determine the amount of reserves in the banking system. Because of this involvement in the Treasury securities market, as well as its walking-distance location near the New York Stock Exchange, the officials at the Federal Reserve Bank of New York are in constant contact with the major financial markets in the United States. In addition, the Federal Reserve Bank of New York houses the foreign exchange desk, which conducts foreign exchange interventions on behalf of the Federal Reserve System and the U.S. Treasury. Its involvement in these financial markets means that the New York Fed is an important source of information on what is happening in U.S. and foreign financial markets, particularly during crisis periods, such as the subprime mortgage meltdown from 2007 to 2009, as well as a liaison between officials in the Federal Reserve System and private participants in the markets.

The third reason for the Federal Reserve Bank of New York's prominence is that it is the only Federal Reserve bank to be a member of the Bank for International Settlements (BIS). Thus the president of the New York Fed, along with the chairperson of the Board of Governors, represents the Federal Reserve System in its regular monthly meetings with other major central bankers at the BIS. This close contact with foreign central bankers and interaction with foreign exchange markets means that the New York Fed has a special role in international relations, both with other central bankers and with private market participants. Adding to its prominence in international circles is that the New York Fed is the repository for more than US$100 billion of the world's gold, an amount greater than the gold at Fort Knox.

Finally, the president of the Federal Reserve Bank of New York is the only permanent member of the FOMC among the Federal Reserve bank presidents, serving as the vice-chair of the committee. Thus they and the chair and vice-chair of the Board of Governors are the three most important officials in the Federal Reserve System.

chairperson of the Board of Governors appointed by a previous administration. Alan Greenspan, for example, was appointed chairperson in 1987 by President Ronald Reagan and was reappointed to another term by another Republican president, George Bush, in 1992. When Bill Clinton, a Democrat, became president in 1993, Greenspan had several years left to his term. Clinton was put under tremendous pressure to reappoint Greenspan when his term expired and did so in 1996 and again in 2000, even though Greenspan is a Republican.[14] George W. Bush, a Republican, then reappointed

[14]Similarly, William McChesney Martin Jr., the chairperson from 1951 to 1970, was appointed by President Truman (Dem.) but was reappointed by Presidents Eisenhower (Rep.), Kennedy (Dem.), and Nixon (Rep.). Also, Paul Volcker, the chairperson from 1979 to 1987, was appointed by President Carter (Dem.) but was reappointed by President Reagan (Rep.).

FYI U.S. Presidential Attacks on the Independence of the Fed

U.S. presidents have often tried to influence Federal Reserve policies by attacking Fed independence. Lyndon Johnson privately made appeals to the then-chair of the Federal Reserve, William McChesney Martin, to not raise interest rates at the December 1965 FOMC meeting and considered trying to fire Martin but was informed that he could not legally fire a Fed chair for just disagreeing with a presidential administration's policies. After the FOMC raised rates at the December FOMC meeting, Johnson was livid and summoned Martin to fly down to the Johnson Ranch in Texas for what Johnson referred to as "a trip to the woodshed." Johnson was not above using physical force to intimidate Martin: The 6-foot 4-inch Johnson actually shoved the shorter Martin up against a wall. Martin stood his ground and told Johnson that the Federal Reserve Act gave the Fed, not the president or Congress, the responsibility for setting interest rates.

Richard Nixon also leaned on his appointed chair of the Federal Reserve, Arthur Burns, to keep interest rates low in the run-up to the 1972 election. Arthur Burns was very close to Nixon, and many economists have argued that Burns gave in to Nixon, dealing a blow to Federal Reserve independence. Subsequently, inflation rose sharply, leading to a period that economists have referred to as the Great Inflation. Ronald Reagan and George H. W. Bush also expressed their unhappiness with Paul Volcker's push to raise the federal funds rate to extremely high levels (20%) in Volcker's successful attempt to control inflation. However, neither of them took overt steps to weaken the independence of the Fed.

When President Clinton came into office, Robert Rubin, who initially had a position in the White House and subsequently became Treasury secretary, convinced Clinton that he should respect the independence of the Fed and not publicly comment on Fed policies. This so-called Rubin doctrine—that the independence of the Fed should be respected by having the president not comment on Fed policies—was then adopted by George W. Bush and later by Barack Obama.

With the presidency of Donald Trump, the Rubin doctrine was totally abandoned. The virulence of the public attacks by Trump on the Fed starting in 2018 were unprecedented. Beginning in July 2018, Trump complained publicly that he was "not thrilled" with the Fed's raising interest rates. At a campaign rally in October 2018, he escalated his attacks on the Fed by saying that the Fed's monetary policy was too tight and that the "Fed has gone crazy." Shortly afterward he gave a telephone interview on Fox News in which he expressed that the Fed was "going loco" when it raised interest rates. Starting in 2019, Trump began to attack Jay Powell personally by telling numerous media outlets that he was unhappy with his selection of Jay Powell and "maybe" regretted appointing Powell to the chair of the Federal Reserve Board. In addition, he accused Powell of being "clueless" and of having "a horrendous lack of vision." The White House also leaked to the press that Trump was discussing firing Powell, but Trump was then advised that he did not have the legal right to do so. Trump also tweeted that the economy would soar "like a rocket" if the Fed would just lower the federal funds rate and engage in large-scale purchases of bonds. When the Fed did not accede to his requests, he tweeted that the Fed was behaving like a "stubborn child" and that its officials were "boneheads."

Chair Powell has vociferously defended the independence of the Federal Reserve and has stated that the Fed has not acceded to pressure from the White House. However, when the Fed started to lower interest rates at the end of July 2019, many economic commentators took the opposite view, that the Federal Reserve was giving in to presidential pressure, which would compromise the independence of the Federal Reserve.

Greenspan in 2004. Presidents have attacked Federal Reserve independence, although it is not clear how successful they have been (see the Inside the Fed box "Presidential Attacks on the Independence of the Fed").

You can see that the Federal Reserve has extraordinary independence for a government agency. Nonetheless, the Fed is not free from political pressures. Indeed, to

understand the Fed's behaviour, we must recognize that public support for the actions of the Federal Reserve plays a very important role in its decisions.[15]

European Central Bank

Until recently, the Federal Reserve had no rivals in terms of importance in the central banking world. However, this situation changed in January 1999 with the full operational start-up of the European Central Bank (ECB) and European System of Central Banks (ESCB), which now conduct monetary policy for countries that are members of the European Monetary Union. These countries, taken together, have a population that exceeds that of the United States and a GDP comparable to that of the United States. The Maastricht Treaty, which established the ECB and ESCB, patterned these institutions after the Federal Reserve, in that central banks for each member country (referred to as *National Central Banks*, or *NCBs*) have a similar role to that of the Federal Reserve Banks. The European Central Bank and the National Central Banks of the countries that have adopted the euro are referred to as the **Eurosystem**. The Eurosystem excludes the central banks of the European Union countries that have not yet adopted the euro. The Eurosystem should not be confused with the European System of Central Banks, the term used to refer to the European Central Bank and the National Central Banks of all 27 European Union (EU) member countries (the United Kingdom withdrew from the EU on January 31, 2020). When all EU member countries adopt the euro, the Eurosystem and the European System of Central Banks will become one entity.

The European Central Bank, which is housed in Frankfurt, Germany, has an Executive Board that is similar in structure to the Board of Governors of the Federal Reserve; it is made up of the president, the vice-president, and four other members, who are appointed to eight-year, nonrenewable terms. The Governing Council, which comprises the Executive Board and the presidents of the National Central Banks, is similar to the FOMC and makes the decisions on monetary policy. While the presidents of the National Central Banks are appointed by their countries' governments, the members of the Executive Board are appointed by a committee consisting of the heads of state of all the countries that are part of the European Monetary Union.

Differences Between the European System of Central Banks and the Federal Reserve System

In the popular press, the European System of Central Banks is usually referred to as the ECB, even though it would be more accurate to refer to it as the Eurosystem, just as it would be more accurate to refer to the Fed as the Federal Reserve System. Although the structure of the Eurosystem is similar to that of the Federal Reserve System, some important differences distinguish the two. First, the budgets of the Federal Reserve Banks are controlled by the Board of Governors, whereas the National Central Banks control their own budgets *and* the budget of the ECB in Frankfurt. The ECB in the Eurosystem therefore has less power than does the Board of Governors in the Federal Reserve System. Second, the monetary operations of the Eurosystem are conducted by the National Central Banks in each country, so monetary operations are not centralized as they are in the Federal Reserve System.

Governing Council

Just as there is a focus on meetings of the FOMC in the United States, there is a similar focus in Europe on meetings of the Governing Council,

[15]An inside view of how the Fed interacts with the public and the politicians can be found in Bob Woodward, *Maestro: Greenspan's Fed and the American Boom* (New York: Simon and Schuster, 2000).

which gathers monthly at the ECB in Frankfurt to make decisions on monetary policy. Currently, 19 countries are members of the European Monetary Union, and the heads of the 19 National Central Banks share 15 votes on the Governing Council; each of the 6 Executive Board members has 1 vote. In contrast to FOMC meetings, which are attended by staff from both the Board of Governors and the individual Federal Reserve Banks, only the 25 members of the Governing Council attend its meetings, with no staff present.

The Governing Council has decided that although its members have the legal right to vote, no formal vote will actually be taken; instead, the Council operates by consensus. One reason the Governing Council has decided not to take votes is because of worries that the casting of individual votes might lead the heads of National Central Banks to support a monetary policy that would be appropriate for their individual countries but not necessarily for the countries of the European Monetary Union as a whole. This problem is less severe for the Federal Reserve: although Federal Reserve bank presidents do live in different regions of the United States, all have the same nationality and are more likely to take a national view in monetary policy decisions rather than a regional view.

Just as the Federal Reserve releases the FOMC's decision on the setting of the policy interest rate (the federal funds rate) immediately after the meeting is over, the ECB does the same after the conclusion of the monthly Governing Council meeting (by announcing the target for a similar short-term interest rate for interbank loans). In addition, immediately after every monetary policy meeting of the Governing Council, the president and vice-president of the ECB hold a press conference in which they take questions from the news media. (A similar press conference is held by the Federal Reserve chair, but it is less frequent, held only four times a year.) The large number of members in the Governing Council presents a particular dilemma. The current size of the Governing Council (25 members, with 21 voting members) is substantially larger than that of the FOMC (19 participants, with 12 voting members). Many commentators have wondered whether the Governing Council is already too unwieldy—a situation that will get considerably worse as more countries join the European Monetary Union. To deal with this potential problem, the Governing Council decided on a complex system of rotation, somewhat like that of the FOMC, in which National Central Banks from the larger countries will vote more often than National Central Banks from the smaller countries.

How Independent Is the ECB? Although the Federal Reserve is a highly independent central bank, the Maastricht Treaty, which established the Eurosystem, has made the latter the most independent central bank in the world. Like the Board of Governors, the members of the Executive Board have long terms (eight years), while heads of National Central Banks are required to serve for at least five years. Like the Fed, the Eurosystem determines its own budget, and the governments of the member countries are not allowed to issue instructions to the ECB. These elements of the Maastricht Treaty make the ECB highly independent.

The Maastricht Treaty specifies that the overriding, long-term goal of the ECB is price stability, which means that the goal for the Eurosystem is more clearly specified than it is for the Federal Reserve System. However, the Maastricht Treaty did not specify exactly what "price stability" means. The Eurosystem has defined the quantitative goal for monetary policy to be an inflation rate slightly less than 2%, so from this perspective, the ECB is slightly less goal-independent than the Fed. The Eurosystem is, however, much more goal-independent than the Federal Reserve System in one

way: The Eurosystem's charter cannot be changed by legislation; it can be changed only by revision of the Maastricht Treaty—a difficult process because *all* signatories to the treaty must agree to accept any proposed change.

Bank of England

Founded in 1694, the Bank of England is the second-oldest central bank (the Riksbank of Sweden is the oldest). The Bank Act of 1946 gave the government statutory authority over the Bank of England. The Court (equivalent to a board of directors) of the Bank of England is made up of the governor and two deputy governors, who are appointed for five-year terms, and 16 non-executive directors, who are appointed for three-year terms.

Until 1997, the Bank of England was the least independent of the central banks examined in this chapter because the decision to raise or lower interest rates resided not within the Bank of England but with the Chancellor of the Exchequer (the equivalent of the Canadian minister of finance). All of this changed when a new Labour government came to power in May 1997 and the Chancellor of the Exchequer, Gordon Brown, made a surprise announcement that the Bank of England would henceforth have the power to set interest rates. However, the Bank was not granted total instrument independence: The government can overrule the Bank and set rates "in extreme economic circumstances" and "for a limited period." Nonetheless, as in Canada, overruling the Bank is likely to be a rare occurrence, because doing so would be highly public and is meant to occur only in highly unusual circumstances and for a limited time.

Because the United Kingdom is not a member of the European Monetary Union, the Bank of England makes its monetary policy decisions independently from the European Central Bank. The power to set interest rates resides with the Monetary Policy Committee, made up of the governor, two deputy governors, two members (normally central bank officials) appointed by the governor after consultation with the chancellor, plus four outside economic experts appointed by the chancellor. (Surprisingly, two of the four outside experts initially appointed to this committee were not British citizens—one was Dutch and the other American—and some later appointments have also not been British citizens, including the previous governor, Mark Carney, who is Canadian.) The inflation target for the Bank of England is set by the Chancellor of the Exchequer, so the Bank of England is also less goal-independent than the Fed.

Bank of Japan

The Bank of Japan (Nippon Ginko) was founded in 1882 during the Meiji Restoration. Monetary policy is determined by the Bank's Policy Board, which is composed of the governor, two vice-governors, and six outside members appointed by the cabinet and approved by the parliament, all of whom serve for five-year terms.

Until recently, the Bank of Japan was not formally independent of the government, with the ultimate power residing with the Ministry of Finance. However, the Bank of Japan Law, which took effect in April 1998 and was the first major change in the relative power of the Bank of Japan in 55 years, changed this situation. In addition to stipulating that the objective of monetary policy is to attain price stability, the law granted greater instrument and goal independence to the Bank of Japan. Before this, the government had two voting members on the Policy Board: one from the Ministry of Finance and the other from the Economic Planning Agency. Now the government may send two representatives from these agencies to board meetings, but they no longer

have voting rights, although they are allowed to request delays in monetary policy decisions. In addition, the Ministry of Finance lost its authority to oversee many operations of the Bank of Japan, particularly the right to dismiss senior officials. However, the Ministry of Finance continues to have control over the part of the Bank's budget that is unrelated to monetary policy. An incident in 2013, in which Japan's Abe government put pressure on the Bank of Japan to adopt a 2% inflation target, against the wishes of the Bank's current governor, who then resigned, suggests that the Bank of Japan's independence is limited.

The Trend Toward Greater Independence

As our survey of the structure and independence of the major central banks indicates, in recent years we have been seeing a remarkable trend toward increasing independence of central banks. It used to be that the Federal Reserve was substantially more independent than almost all other central banks, with the exception of those in Germany and Switzerland. Now the newly established European Central Bank is far more independent than the Fed, and greater independence has been granted to central banks like the Bank of England and the Bank of Japan, putting them more on a par with the Fed, as well as to central banks in such diverse countries as New Zealand, Sweden, and the euro nations. Both theory and experience suggest that more independent central banks produce better monetary policy, thus providing an impetus for this trend.

SUMMARY

1. The Bank of Canada was created by an act of Parliament in 1934 and began operations on March 11, 1935. Initially it was privately owned but became a Crown corporation in 1938.
2. The overall responsibility for the operation of the Bank of Canada rests with a Board of Directors, consisting of the governor, the senior deputy governor, the deputy minister of finance, and 12 outside directors. The Bank's governor is the chief executive officer and chair of the Board of Directors.
3. Although on paper the Bank of Canada is an arm of the government, in practice the Bank has more independence than the Bank of Canada Act suggests.
4. The Bank of Canada is more independent than most agencies of the Canadian government, but it is still subject to political pressures. The theory of bureaucratic behaviour indicates that one factor driving the Bank's behaviour is its attempt to increase its power and prestige. This view explains many of the Bank's actions, although the agency may also try to act in the public interest.
5. The case for an independent Bank of Canada rests on the view that curtailing the Bank's independence and subjecting it to more political pressures would impart an inflationary bias to monetary policy. An independent Bank of Canada can afford to take the long view and not respond to short-run problems that will result in expansionary monetary policy and a political business cycle. The case against an independent Bank of Canada holds that it is undemocratic to have monetary policy (so important to the public) controlled by an elite group that is not accountable to the public. An independent Bank of Canada also makes the coordination of monetary and fiscal policy difficult.
6. The U.S. Federal Reserve System was created in 1913 to lessen the frequency of bank panics. Because of public hostility to central banks and the centralization of power, the Federal Reserve System was created with many checks and balances to diffuse power.
7. The Federal Reserve System consists of 12 regional Federal Reserve banks, around 3000 member commercial banks, the Board of Governors of the Federal Reserve System, the Federal Open Market Committee, and the Federal Advisory Council.
8. Although on paper the Federal Reserve appears to be decentralized, in practice it has come to function as a unified central bank controlled by the Board of Governors, especially the board's chair.

9. The European System of Central Banks has a similar structure to that of the Federal Reserve System, with each member country having a National Central Bank, and an Executive Board of the European Central Bank (located in Frankfurt, Germany). The Governing Council, which is made up of the six members of the Executive Board (including the president of the European Central Bank) and the presidents of the National Central Banks, makes the decisions on monetary policy. The Eurosystem, which was established under the terms of the Maastricht Treaty, is even more independent than the Federal Reserve System because its charter cannot be changed by legislation. Indeed, it is the most independent central bank in the world.
10. There has been a remarkable trend toward increasing independence of central banks throughout the world. Greater independence has been granted to central banks such as the Bank of England and the Bank of Japan in recent years, as well as to other central banks in such diverse countries as New Zealand and Sweden. Both theory and experience suggest that more independent central banks produce better monetary policy.

MyLab Economics Make the grade with MyLab Economics: Select end-of-chapter questions can be found on MyLab Economics. You can practise them as often as you want, and most feature step-by-step guided instructions to help you find the right answer.

QUESTIONS

1. What political realities might explain the creation of the Bank of Canada in 1934?
2. In what ways can the government influence the conduct of monetary policy?
3. Who is responsible for monetary policy in Canada?
4. Do you think that the seven-year renewable term for the governor of the Bank of Canada effectively insulates the Bank from political pressure?
5. How did the "Coyne Affair" motivate the current system of joint responsibility for monetary policy?
6. Over time, which entities have gained power in the conduct of monetary policy in Canada and which have lost power? Why do you think this has happened?
7. "The strongest argument for an independent Bank of Canada rests on the view that subjecting the Bank of Canada to more political pressures would impart an inflationary bias to monetary policy." Is this statement true, false, or uncertain? Explain your answer.
8. The Bank of Canada is the most independent of all Canadian government agencies. What is the main difference between it and other government agencies that explains its greater independence?
9. How can the government directly influence the Bank of Canada?
10. How is the responsibility for monetary policy shared in the Eurosystem?
11. "The theory of bureaucratic behaviour indicates that the Bank of Canada never operates in the public interest." Is this statement true, false, or uncertain? Explain your answer.
12. Why might eliminating the Bank of Canada's independence lead to a more pronounced political business cycle?
13. "The independence of the Bank of Canada leaves it completely unaccountable for its actions." Is this statement true, false, or uncertain? Explain your answer.
14. "The independence of the Bank of Canada means that it takes the long view and not the short view." Is this statement true, false, or uncertain? Explain your answer.
15. Why did Canada show little interest in the establishment of a central bank during the first 60 or so years of Confederation?
16. What is the primary tool that the U.S. Congress uses to exercise control over the Fed?
17. Should the Federal Reserve be subject to periodic auditing of its policies, procedures, and finances? Why or why not?
18. The Fed promotes secrecy by not releasing the minutes of FOMC meetings to Congress or the public immediately. Discuss the arguments for and against this policy.
19. Which is more independent, the Federal Reserve or the European Central Bank? Why?
20. Why did the Bank of England up until 1997 have a low degree of independence?

DATA ANALYSIS PROBLEMS

1. Go to the CANSIM database and find data on the unemployment rate in New Brunswick (series V2063760), Ontario (series V2063949), and British Columbia (series V2064705). For the most recent month of data available, determine which province had the highest and which had the lowest unemployment rate.

2. Go to the Bank of Canada's website and find data on the overnight interest rate target. When was the last time the overnight interest rate target was changed? Did the rate increase or decrease?

15 The Money Supply Process

Learning Objectives

15.1 List and describe the "three players" that influence the money supply.

15.2 Classify the factors affecting the Bank of Canada's assets and liabilities.

15.3 Identify the factors that affect the monetary base, and discuss their effects on the Bank of Canada's balance sheet.

15.4 Explain and illustrate the deposit creation process through T-accounts.

15.5 List the factors that affect the money supply.

15.6 Summarize how the "three players" can influence the money supply.

15.7 Calculate and interpret changes in the money multiplier.

Preview

As we saw in Chapter 5 and will see in later chapters on monetary theory, movements in the money supply affect interest rates and inflation and thus affect us all. Because of its far-reaching effects on economic activity, it is important to understand how the money supply is determined. Who controls it? What causes it to change? How might control of it be improved? In this chapter, we start to answer these questions by providing a detailed description of the *money supply process*, the mechanism that determines the level of the money supply.

Because deposits at banks are by far the largest component of the money supply, learning how these deposits are created is the first step in understanding the money supply process. This chapter provides an overview of how the banking system creates deposits and describes the basic principles of the money supply, concepts that will form the foundation for the material presented in later chapters.

15.1 THREE PLAYERS IN THE MONEY SUPPLY PROCESS

LO 15.1 List and describe the "three players" that influence the money supply.

The "cast of characters" in the money supply story is as follows:

1. The *central bank*—the government agency that oversees the banking system and is responsible for the conduct of monetary policy; in Canada, the Bank of Canada
2. *Banks* (depository institutions)—the financial intermediaries that accept deposits from individuals and institutions and make loans: commercial banks, trust and loan companies, and credit unions
3. *Depositors*—individuals and institutions that hold deposits in banks

Of the three players, the central bank—the Bank of Canada—is the most important. The Bank's conduct of monetary policy involves actions that affect its balance sheet (holdings of assets and liabilities), to which we turn now.

15.2 THE BANK OF CANADA'S BALANCE SHEET

LO 15.2 Classify the factors affecting the Bank of Canada's assets and liabilities.

The operation of the Bank of Canada and its monetary policy involve actions that affect its balance sheet, or its holdings of assets and liabilities. Here we discuss a simplified balance sheet that includes just four items that are essential to our understanding of the money supply process.[1]

Bank of Canada

Assets	Liabilities
Securities	Currency in circulation
Loans to financial institutions	Reserves

Liabilities

The two liabilities on the balance sheet, currency in circulation and reserves, are often referred to as the *monetary liabilities* of the Bank of Canada. They are an important part of the money supply story, because increases in either or both will lead to an increase in the money supply (everything else held constant). The sum of the Bank's monetary liabilities (currency in circulation and reserves) and the Canadian Mint's monetary liabilities, coins in circulation, is called the **monetary base** (also called *high-powered money*). When discussing the monetary base, we will focus only on the monetary liabilities of the Bank of Canada, because the monetary liabilities of the Canadian Mint account for a very small fraction of the base.

1. *Notes in circulation.* The Bank of Canada issues notes (those blue, purple, green, red, and brown pieces of paper in your wallet that say "Bank of Canada"). The Bank of Canada notes in circulation is the amount of these notes that is in the hands of the public and the depository institutions. Coins issued by the Canadian Mint are not a liability of the Bank of Canada. The coins and Bank of Canada notes that we use in Canada today are collectively known as *currency*.

   Bank of Canada notes are IOUs from the Bank to the bearer and are also liabilities, but unlike most liabilities, they promise to pay back the bearer solely with Bank of Canada notes; that is, they pay off IOUs with other IOUs. Accordingly, if you bring a $100 bill to the Bank of Canada and demand payment, you will receive two $50s, five $20s, ten $10s, twenty $5 bills, or some other combination of bills that adds to $100.

   People are more willing to accept IOUs from the Bank of Canada than from you or me because Bank of Canada notes are a recognized medium of exchange; that is, they are accepted as a means of payment and so function as *money*. Unfortunately,

[1]A detailed discussion of the Bank of Canada's balance sheet and the factors that affect the monetary base can be found in Appendix 15.1 to this chapter, which you can find online.

neither you nor I can convince people that our IOUs are worth anything more than the paper they are written on.[2]

2. *Reserves.* All banks that participate in the Large Value Transfer System, to be discussed in detail in Chapter 16, have an account at the Bank of Canada in which they hold deposits (also called **settlement balances**). **Reserves** consist of settlement balances at the Bank of Canada plus currency that is physically held by banks (called *vault cash*, because it is held in bank vaults, cash tills, and automated banking machines). Reserves are assets for the banks but liabilities for the Bank of Canada, because the banks can demand payment on them at any time, and the Bank of Canada is required to satisfy its obligation by paying Bank of Canada notes. As you will see, an increase in reserves leads to an increase in the level of deposits and hence in the money supply.

   As already noted in Chapter 12, Canadian banks are not required to hold reserves. Banks, however, hold reserves in order to manage their own short-term liquidity requirements and to respond to predictable clearing drains and across-the-counter and automated banking machine drains. We call these reserves prudential or **desired reserves**. For example, banks might desire that for every dollar of deposits, a certain fraction (say, 5 cents) must be held as reserves. This fraction (5%) is called the **desired reserve ratio**. Reserves in excess of the desired amounts are called **excess reserves**.

Assets

The two assets on the Bank of Canada's balance sheet are important for two reasons. First, changes in the asset items lead to changes in reserves and the monetary base and consequently to changes in the money supply. Second, because these assets (government securities and Bank of Canada loans) earn higher interest rates than the liabilities do (currency in circulation, which pays no interest, and reserves), the Bank of Canada makes billions of dollars every year—its assets earn income, and its liabilities cost practically nothing. Although it returns most of its earnings to the federal government, the Bank of Canada does spend some of it on "worthy causes," such as supporting economic research.

1. *Securities.* This category of assets covers the Bank of Canada's holdings of securities issued by the Canadian government and, in unusual circumstances (as discussed in Chapter 16), other securities. As you will see, the primary way the Bank of Canada provides reserves to the banking system is by purchasing securities, thereby increasing its holdings of these assets. An increase in government or other securities held by the Bank of Canada leads to an increase in the money supply.
2. *Loans to financial institutions.* The second way the Bank of Canada can provide reserves to the banking system is by making loans (equivalently, advances) to banks and other financial institutions. The loans taken out by these institutions

[2]The notes item on our balance sheet refers only to notes in circulation—that is, the amount in the hands of the public. Notes that have been printed are not automatically a liability of the Bank. For example, consider the importance of having $1 million of your own IOUs printed up. You give out $100 worth to other people and keep the other $999 900 in your pocket. The $999 900 of IOUs does not make you richer or poorer and does not affect your indebtedness. You care only about the $100 of liabilities from the $100 of circulated IOUs. The same reasoning applies for the Bank of Canada in regard to its notes.

For similar reasons, the currency component of the money supply, no matter how it is defined, includes only currency in circulation. It does not include any additional currency that is not yet in the hands of the public. The fact that currency has been printed but is not circulating means that it is not anyone's asset or liability and thus cannot affect anyone's behaviour. Therefore, it makes sense not to include it in the money supply.

are referred to as *borrowings from the Bank of Canada* or, alternatively, as *borrowed reserves*. These loans appear as a liability on financial institutions' balance sheets. An increase in loans to financial institutions can also be the source of an increase in the money supply. During normal times, the Bank of Canada makes loans only to banking institutions, and the interest rate charged to banks for these loans is called the **bank rate**. (As we will discuss in Chapter 16, during the 2008 financial crisis, the Bank of Canada made loans to other financial institutions.)

15.3 CONTROL OF THE MONETARY BASE

LO 15.3 Identify the factors that affect the monetary base, and discuss their effects on the Bank of Canada's balance sheet.

The monetary base equals currency in circulation C plus the total reserves in banking system R.[3] The monetary base MB can be expressed as

$$MB = C + R$$

The Bank of Canada exercises control over the monetary base through its purchases or sales of securities in the open market, called **open market operations**, and through its extension of discount loans to banks.

Bank of Canada Open Market Operations

The primary way in which the Bank of Canada causes changes in the monetary base is through its open market operations. A purchase of bonds by the Bank of Canada is called an **open market purchase**, and a sale of bonds by the Bank of Canada is called an **open market sale**. Bank of Canada purchases and sales of bonds are always done through **primary dealers**, government securities dealers who operate out of private banking institutions.

Open Market Purchase

Suppose the Bank of Canada purchases $100 million of bonds from a primary dealer. To understand the consequences of this transaction, we look at *T-accounts*, which list only the changes that occur in balance sheet items, starting from the initial balance sheet position.

When the primary dealer sells the $100 million of bonds to the Bank of Canada, the Bank of Canada adds $100 million to the dealer's deposit account at the Bank of Canada so that reserves in the banking system go up by $100 million. The banking system's T-account after this transaction is

Banking System

Assets		Liabilities
Securities	−$100 m	
Reserves	+$100 m	

[3]Here, currency in circulation includes both Bank of Canada notes and Canadian Mint coins.

The effects on the Bank of Canada's balance sheet are shown next. The balance sheet shows an increase of $100 million of securities in its assets column, along with an increase of $100 million of reserves in its liabilities column:

Bank of Canada

Assets		Liabilities	
Securities	+$100 m	Reserves	+$100 m

As you can see, the Bank of Canada's open market purchase of $100 million causes an expansion of reserves in the banking system by an equal amount. Another way of seeing this is to recognize that open market purchases of bonds expand reserves because the central bank pays for the bonds with reserves. Because the monetary base equals currency plus reserves, an open market purchase increases the monetary base by an amount equal to the amount of the purchase.

Open Market Sale Similar reasoning indicates that if the Bank of Canada conducts an open market sale of $100 million of bonds to a primary dealer, the Bank of Canada deducts $100 million from the dealer's deposit account, so the Bank of Canada's reserves (liabilities) fall by $100 million (and the monetary base falls by the same amount). The T-account is now

Bank of Canada

Assets		Liabilities	
Securities	−$100 m	Reserves	−$100 m

Although open market operations are the most important monetary policy tool for most central banks around the world, in 1994 the Bank of Canada stopped conducting open market operations in Government of Canada bills and bonds. Since then, the Bank of Canada's most common open market operations involve repurchase transactions, either SPRAs or SRAs. More recently, and in response to the global financial crisis, the Bank of Canada introduced term PRAs, which are similar to special PRAs. As we will discuss in detail in Chapter 16, the Bank of Canada is currently conducting repurchase transactions to reinforce its operating target—the midpoint of the operating band for the overnight interest rate—and to address liquidity issues at times of financial instability. The Bank of Canada, however, neutralizes the effect on settlement balances of SPRAs and SRAs so that at the end of the day there is no change in the level of settlement balances in the banking system.

Shifts from Deposits into Currency

Even when the Bank of Canada does not conduct open market operations, a shift from deposits to currency will affect the reserves in the banking system. However, such a shift will have no effect on the monetary base. This tells us that the Bank of Canada has more control over the monetary base than over reserves.

Let's suppose that during the Christmas season, the public wants to hold more currency to buy gifts and so withdraws $100 million in cash. The effect on the T-account of the nonbank public is

Nonbank Public

Assets		Liabilities
Chequable deposits	−$100 m	
Currency	+$100 m	

The banking system loses $100 million of deposits and hence $100 million of reserves:

Banking System

Assets		Liabilities	
Reserves	−$100 m	Chequable deposits	−$100 m

For the Bank of Canada, the public's action means that $100 million of additional currency is circulating in the hands of the public, while reserves in the banking system have fallen by $100 million. The Bank of Canada's T-account is

Bank of Canada

Assets	Liabilities	
	Currency in circulation	+$100 m
	Reserves	−$100 m

The net effect on the monetary liabilities of the Bank of Canada is a wash; the monetary base is unaffected by the public's increased desire for cash. But reserves *are* affected. Random fluctuations of reserves can occur as a result of random shifts into currency and out of deposits, and vice versa. The same is not true for the monetary base, making it a more stable variable and more controllable by the Bank of Canada.

Loans to Financial Institutions

In this chapter so far, we have seen how changes in the monetary base occur as a result of open market operations. However, the monetary base is also affected when the Bank of Canada makes a loan to a financial institution. When the Bank of Canada makes a $100 million loan to the First National Bank, the bank is credited with $100 million of

reserves from the proceeds of the loan. The effects on the balance sheets of the banking system and the Bank of Canada are illustrated by the following T-accounts:

Banking System				Bank of Canada			
Assets		**Liabilities**		**Assets**		**Liabilities**	
Reserves	+$100 m	Loans (borrowings from the Bank of Canada)	+$100 m	Loans (borrowings from the Bank of Canada)	+$100 m	Reserves	+$100 m

The monetary liabilities of the Bank of Canada have now increased by $100 million, and the monetary base, too, has increased by this amount. However, if a bank pays off a loan from the Bank of Canada, thereby reducing its borrowings from the Bank of Canada by $100 million, the T-accounts of the banking system and the Bank of Canada are as follows:

Banking System				Bank of Canada			
Assets		**Liabilities**		**Assets**		**Liabilities**	
Reserves	−$100 m	Loans (borrowings from the Bank of Canada)	−$100 m	Loans (borrowings from the Bank of Canada)	−$100 m	Reserves	−$100 m

The net effect on the monetary liabilities of the Bank of Canada, and hence on the monetary base, is a reduction of $100 million. We see that the monetary base changes in a one-to-one ratio with the change in the borrowings from the Bank of Canada.

Other Factors That Affect the Monetary Base

So far in this chapter, it seems as though the Bank of Canada has complete control of the monetary base through its open market operations and loans to financial institutions. However, the world is a little bit more complicated for the Bank of Canada. Two important items that affect the monetary base, but are not controlled by the Bank of Canada, are *float* and *government deposits at the Bank of Canada*. When the Bank of Canada clears cheques for banks, it often credits the amount of the cheque to a bank that has deposited it (increases the bank's reserves) before it debits (decreases the reserves of) the bank on which the cheque is drawn. The resulting temporary net increase in the total amount of reserves in the banking system (and hence in the monetary base) caused by the Bank of Canada's cheque-clearing process is called **float**. When the government moves deposits from commercial banks to its account at the Bank of Canada,

leading to an increase in *government deposits at the Bank of Canada*, it causes a deposit outflow at these banks, such as that shown in Chapter 12, and thus causes reserves in the banking system and the monetary base to decrease. Thus *float* (affected by random events such as the weather, which influences how quickly cheques are presented for payment) and *goverment deposits at the Bank of Canada* (determined by the federal government's actions) both affect the monetary base but are not controlled by the Bank of Canada at all. Bank of Canada intervention in the foreign exchange market also affects the monetary base.

Overview of the Bank of Canada's Ability to Control the Monetary Base

Our discussion above indicates that two primary features determine the monetary base: open market operations and lending to financial institutions. Whereas the amount of open market purchases or sales is completely controlled by the Bank of Canada placing orders with dealers in bond markets, the central bank cannot unilaterally determine, and therefore cannot perfectly predict, the amount of borrowings from the Bank of Canada. The Bank of Canada sets the bank rate (interest rate on loans to banks) and then banks make decisions about whether to borrow. The amount of lending, though influenced by the Bank of Canada's setting of the bank rate, is not completely controlled by the Bank of Canada; banks' decisions play a role, too.

Therefore, we might want to split the monetary base into two components: one that the Bank of Canada can control completely and another that is less tightly controlled. The less tightly controlled component is the amount of the base that is created by loans from the Bank of Canada. The remainder of the base (called the **nonborrowed monetary base**) is under the Bank of Canada's control because it results primarily from open market operations.[4] The nonborrowed monetary base is formally defined as the monetary base minus borrowings from the Bank of Canada, which are referred to as **borrowed reserves**:

$$MB_n = MB - BR$$

where MB_n = nonborrowed monetary base
MB = monetary base
BR = borrowed reserves from the Bank of Canada

Factors not controlled at all by the Bank of Canada (for example, float and government deposits with the Bank of Canada) undergo substantial short-run variations and can be important sources of fluctuations in the monetary base over time periods as short as a week. However, these fluctuations are usually predictable and so can be offset through open market operations. ***Although float and government deposits with the Bank of Canada undergo substantial short-run fluctuations, which complicate control of the monetary base, they do not prevent the Bank of Canada from accurately controlling it.***

[4]Actually, other items on the Bank of Canada's balance sheet (discussed in the first online appendix to Chapter 15) affect the magnitude of the nonborrowed monetary base. Because their effects on the nonborrowed base relative to open market operations are both small and predictable, these other items do not present the Bank of Canada with difficulties in controlling the nonborrowed base.

15.4 MULTIPLE DEPOSIT CREATION: A SIMPLE MODEL

LO 15.4 Explain and illustrate the deposit creation process through T-accounts.

With our understanding of how the Bank of Canada controls the monetary base and how banks operate (Chapter 12), we now have the tools necessary to explain how deposits are created. When the Bank of Canada supplies the banking system with $1 of additional reserves, deposits increase by a multiple of this amount—a process called **multiple deposit creation**.

Deposit Creation: The Single Bank

Suppose the $100 million open market purchase described earlier was conducted with the First National Bank. After the Bank of Canada has bought the $100 million in bonds from the First National Bank, the bank finds that it has an increase in reserves of $100 million. To analyze what the bank will do with these additional reserves, assume that the bank does not want to hold excess reserves because it earns little interest on them. We begin the analysis with the following T-account:

First National Bank

Assets		Liabilities
Securities	−$100 m	
Reserves	+$100 m	

Because the bank has no increase in its chequable deposits, desired reserves remain the same, and the bank finds that its additional $100 million of reserves means that its excess reserves have risen by $100 million. Let's say the bank decides to make a loan equal in amount to the $100 million rise in excess reserves. When the bank makes the loan, it sets up a chequing account for the borrower and puts the proceeds of the loan into this account. In this way, the bank alters its balance sheet by increasing its liabilities with $100 million of chequable deposits and at the same time increasing its assets with the $100 million loan. The resulting T-account looks like this:

First National Bank

Assets		Liabilities	
Securities	−$100 m	Chequable deposits	+$100 m
Reserves	+$100 m		
Loans	+$100 m		

The bank has created chequable deposits by its act of lending. Because chequable deposits are part of the money supply, the bank's act of lending has, in fact, created money.

In its current balance sheet position, the First National Bank still has excess reserves and so might want to make additional loans. However, these reserves will not stay at the bank for very long. The borrowers took out loans not to leave $100 million sitting idle in a chequing account at the First National Bank but to purchase goods and services from other individuals and corporations. When the borrowers make these purchases by writing cheques, the cheques will be deposited at other banks, and the $100 million of reserves will leave the First National Bank. ***As a result, a bank cannot safely make a loan for an amount greater than the excess reserves that it has before it makes the loan.***

The final T-account of the First National Bank is

First National Bank

Assets		Liabilities
Securities	−$100 m	
Loans	+$100 m	

The increase in reserves of $100 million has been converted into additional loans of $100 million at the First National Bank, plus an additional $100 million of deposits that have made their way to other banks. (All the cheques written on accounts at the First National Bank are deposited in banks rather than converted into cash, because we are assuming that the public does not want to hold any additional currency.) Now let's see what happens to these deposits at the other banks.

Deposit Creation: The Banking System

To simplify the analysis, let's assume that the $100 million of deposits created by First National Bank's loan is deposited at Bank A and that this bank and all other banks hold no excess reserves. Bank A's T-account becomes

Bank A

Assets		Liabilities	
Reserves	+$100 m	Chequable deposits	+$100 m

If the desired reserve ratio is 10%, this bank will now find itself with a $10 million increase in desired reserves, leaving it $90 million of excess reserves. Because Bank A (like the First National Bank) does not want to hold on to excess reserves, it will make loans for the entire amount. Its loans and chequable deposits will then increase by $90 million, but when the borrowers spend the $90 million of chequable deposits, they and the reserves at Bank A will fall back down by this same amount. The net result is that Bank A's T-account will look like this:

Bank A

Assets		Liabilities	
Reserves	+$10 m	Chequable deposits	+$100 m
Loans	+$90 m		

If the money spent by the borrowers to whom Bank A lent the $90 million is deposited in another bank, such as Bank B, the T-account for Bank B will be

Bank B

Assets		Liabilities	
Reserves	+$90 m	Chequable deposits	+$90 m

The chequable deposits in the banking system have risen by another $90 million, for a total increase of $190 million ($100 million at Bank A plus $90 million at Bank B). In fact, the distinction between Bank A and Bank B is not necessary to obtain the same result on the overall expansion of deposits. If the borrowers from Bank A write cheques to someone who deposits them at Bank A, the same change in deposits occurs. The T-accounts for Bank B would just apply to Bank A, and its chequable deposits would increase by the total amount of $190 million.

Bank B will want to modify its balance sheet further. It must keep 10% of $90 million ($9 million) as desired reserves and has 90% of $90 million ($81 million) in excess reserves and so can make loans of this amount. Bank B will make loans totalling $81 million to borrowers, who spend the proceeds from the loans. Bank B's T-account will be

Bank B

Assets		Liabilities	
Reserves	+$ 9 m	Chequable deposits	+$90 m
Loans	+$81 m		

The $81 million spent by the borrowers from Bank B will be deposited in another bank (Bank C). Consequently, from the initial $100 million increase of reserves in the banking system, the total increase of chequable deposits in the system so far is $271 million (= $100 m + $90 m + $81 m).

Following the same reasoning, if all banks make loans for the full amount of their excess reserves, further increments in chequable deposits will continue (at Banks C, D, E, and so on), as depicted in Table 15-1. Therefore, the total increase in deposits from the initial $100 increase in reserves will be $1000 million: The increase is tenfold, the reciprocal of the 10% (0.10) desired reserve ratio.

TABLE 15-1 Creation of Deposits (Assuming a 10% Desired Reserve Ratio and a $100 Million Increase in Reserves)

Bank	Increase in Deposits ($)	Increase in Loans ($)	Increase in Reserves ($)
First National	0.00	100.00 m	0.00
A	100.00 m	90.00 m	10.00 m
B	90.00 m	81.00 m	9.00 m
C	81.00 m	72.90 m	8.10 m
D	72.90 m	65.61 m	7.29 m
E	65.61 m	59.05 m	6.56 m
F	59.05 m	53.14 m	5.91 m
.	.	.	.
.	.	.	.
.	.	.	.
.	.	.	.
Total for all banks	1 000.00 m	1 000.00 m	100.00 m

If the banks choose to invest their excess reserves in securities, the result is the same. If Bank A had taken its excess reserves and purchased securities instead of making loans, its T-account would have looked like this:

Bank A

Assets		Liabilities	
Reserves	+$10 m	Chequable deposits	+$100 m
Securities	+$90 m		

When the bank buys $90 million of securities, it writes $90 million in cheques to the sellers of the securities, who in turn deposit the $90 million at a bank such as Bank B. Bank B's chequable deposits increase by $90 million, and the deposit expansion process is the same as before. ***Whether a bank chooses to use its excess reserves to make loans or to purchase securities, the effect on deposit expansion is the same.***

You can now see the difference in deposit creation for a single bank versus the banking system as a whole. Because a single bank can create deposits equal only to the amount of its excess reserves, it cannot by itself generate multiple deposit expansion. A single bank cannot make loans greater in amount than its excess reserves, because the bank will lose these reserves as the deposits created by the loan find their way to other banks. However, the banking system as a whole can generate a multiple expansion of deposits, because when a bank loses its excess reserves, these reserves do not leave the banking system, even though they are lost to the individual bank. So as each bank makes a loan and creates deposits, the reserves find their way to another bank, which uses them to make additional loans and create additional deposits. As you have seen, this process continues until the initial increase in reserves results in a multiple increase in deposits.

The multiple increase in deposits generated from an increase in the banking system's reserves is called the **simple deposit multiplier**.[5] In our example, with a 10% desired reserve ratio, the simple deposit multiplier is 10. More generally, the simple deposit multiplier equals the reciprocal of the desired reserve ratio, expressed as a fraction (for example, $10 = 1/0.10$). So the formula for the multiple expansion of deposits can be written as follows:

$$\Delta D = \frac{1}{r_d} \times \Delta R \tag{1}$$

where ΔD = change in total chequable deposits in the banking system
r_d = desired reserve ratio (0.10 in the example)
ΔR = change in reserves for the banking system ($100 million in the example)

Deriving the Formula for Multiple Deposit Creation

The formula for the multiple creation of deposits can be derived directly using algebra. We obtain the same result for the relationship between a change in deposits and a change in reserves.

Our assumption that banks do not hold on to any excess reserves means that the total amount of desired reserves in the banking system DR will equal the total reserves in the banking system R:

$$DR = R$$

The total amount of desired reserves equals the desired reserve ratio r_d times the total amount of chequable deposits D:

$$DR = r_d \times D$$

Substituting $r_d \times D$ for DR in the first equation,

$$r_d \times D = R$$

and dividing both sides of the preceding equation by rr gives

$$D = \frac{1}{r_d} \times R$$

Taking the change in both sides of this equation and using delta to indicate a change gives

$$\Delta D = \frac{1}{r_d} \times \Delta R$$

which is the same formula for deposit creation given in Equation 1.[6]

[5]This multiplier should not be confused with the Keynesian multiplier, which is derived through a similar step-by-step analysis. That multiplier relates an increase in income to an increase in investment, whereas the simple deposit multiplier relates an increase in deposits to an increase in reserves.

[6]A formal derivation of this formula follows. Using the reasoning in the text, the change in chequable deposits is $\$100 (= \Delta R \times 1)$ plus $\$90 [= \Delta R \times (1 - r_d)]$ plus $\$81 [= \Delta R \times (1 - r_d)^2]$ and so on, which can be rewritten as

$$\Delta D = \Delta R \times [1 + (1 - r_d) + (1 - r_d)^2 + (1 - r_d)^3 + \cdots]$$

Using the formula for the sum of an infinite series found in footnote 3 of Chapter 4, this equation can be rewritten as

$$\Delta D = \Delta R \times \frac{1}{1 - (1 - r_d)} = \frac{1}{r_d} \times \Delta R$$

This derivation provides us with another way of looking at the multiple creation of deposits, because it forces us to examine the banking system as a whole rather than one bank at a time. For the banking system as a whole, deposit creation (or contraction) will stop only when excess reserves in the banking system are zero; that is, the banking system will be in equilibrium when the total amount of desired reserves equals the total amount of reserves, as seen in the equation $DR = R$. When $r_d \times D$ is substituted for DR, the resulting equation $r_d \times D = R$ tells us how high chequable deposits must be for desired reserves to equal total reserves. Accordingly, a given level of reserves in the banking system determines the level of chequable deposits when the banking system is in equilibrium (when $ER = 0$); put another way, the given level of reserves supports a given level of chequable deposits.

In our example, the desired reserve ratio is 10%. If reserves increase by $100 million, chequable deposits must rise by $1000 million for total desired reserves also to increase by $100 million. If the increase in chequable deposits is less than this—say, $900 million—then the increase in desired reserves of $90 million remains below the $100 million increase in reserves, so excess reserves still exist somewhere in the banking system. The banks holding the excess reserves will now make additional loans, thereby creating new deposits; this process will continue until all reserves in the system are used up, which occurs when chequable deposits rise by $1000 million.

We can also see this by looking at the resulting T-account of the banking system as a whole (including the First National Bank):

Banking System

Assets		Liabilities	
Securities	−$ 100 m	Chequable deposits	+$1000 m
Reserves	+$ 100 m		
Loans	+$ 1000 m		

The procedure of eliminating excess reserves by loaning them out continues until the banking system (First National Bank and Banks A, B, C, D, and so on) has made $1000 million of loans and created $1000 million of deposits. In this way, $100 million of reserves supports $1000 million (10 times the quantity) of deposits.

Critique of the Simple Model

Our model of multiple deposit creation seems to indicate that the Bank of Canada is able to exercise complete control over the level of chequable deposits by setting the level of reserves. The actual creation of deposits is much less mechanical than the simple model indicates. If proceeds from Bank A's $90 million loan are not deposited but are kept in currency, nothing is deposited in Bank B and the deposit creation process ceases. The total increase in the money supply is now the $90 million increase in currency plus the initial $100 million of deposits created by First National Bank's loans, which were deposited at Bank A, for a total of only $190 million—considerably less than the $1000 million we calculated using the simple model above. In other words, currency does not lead to multiple deposit expansion, whereas deposits do. Thus, if some proceeds from loans are not deposited in banks but

instead are used to raise the holdings of currency, less multiple expansion occurs overall, and the money supply does not increase by the amount predicted by our simple model of multiple deposit creation.

Another situation ignored in our model is one in which banks do not make loans or buy securities in the full amount of their excess reserves. If Bank A decides to hold on to all $90 million of its excess reserves, no deposits will be made in Bank B, and this will stop the deposit creation process. The total increase in deposits will be only $100 million, not the $1000 million increase in our example. Hence, if banks choose to hold on to all or some of their excess reserves, the full expansion of deposits predicted by the simple model of multiple deposit creation again does not occur.

Our examples indicate that the Bank of Canada is not the only player whose behaviour influences the level of deposits and therefore the money supply. Depositors' decisions regarding how much currency to hold and banks' decisions regarding the amount of reserves to hold can also cause the money supply to change.

15.5 FACTORS THAT DETERMINE THE MONEY SUPPLY

LO 15.5 List the factors that affect the money supply.

Our critique of the simple model shows how we can expand on it to discuss all the factors that affect the money supply. Let's look at changes in each factor in turn, holding all other factors constant.

Changes in the Nonborrowed Monetary Base, MB_n

As shown earlier in the chapter, the Bank of Canada's open market purchases increase the nonborrowed monetary base, and its open market sales decrease it. Holding all other variables constant, an increase in MB_n arising from an open market purchase raises the amount of the monetary base and reserves so that multiple deposit creation occurs and the money supply increases. Similarly, an open market sale that reduces MB_n shrinks the amount of the monetary base and reserves, thereby causing a multiple contraction of deposits and a decrease in the money supply. We have the following result: ***The money supply is positively related to the nonborrowed monetary base MB_n.***

Changes in Borrowed Reserves, *BR*, from the Bank of Canada

An increase in loans from the Bank of Canada provides additional borrowed reserves and thereby increases the amount of the monetary base and reserves so that multiple deposit creation occurs and the money supply expands. If banks reduce the level of their loans from the Bank of Canada, all other variables held constant, the monetary base and amount of reserves fall, and the money supply decreases. The result is this: ***The money supply is positively related to the level of borrowed reserves, BR, from the Bank of Canada.***

Changes in the Desired Reserve Ratio, r_d

If the desired reserve ratio on chequable deposits increases while all other variables, such as the monetary base, stay the same, we have seen that multiple deposit expansion is reduced and hence the money supply falls. If, in contrast,

the desired reserve ratio falls, multiple deposit expansion is higher and the money supply rises.

We now have the following result: ***The money supply is negatively related to the desired reserve ratio r_d.***

Changes in Excess Reserves

When banks increase their holdings of excess reserves, those reserves are no longer being used to make loans, causing multiple deposit creation to stop dead in its tracks, resulting in less expansion of the money supply. If, however, banks choose to hold fewer excess reserves, loans and multiple deposit creation increase, and the money supply rises. ***The money supply is negatively related to the amount of excess reserves.***

Recall from Chapter 12 that the primary benefit to a bank of holding excess reserves is that they provide insurance against losses due to deposit outflows; that is, they enable the bank experiencing deposit outflows to escape the costs of calling in loans, selling securities, borrowing from the Bank of Canada or other corporations, or bank failure. If banks fear that deposit outflows are likely to increase (that is, if expected deposit outflows increase), they will seek more insurance against this possibility, and excess reserves will rise.

Changes in Currency Holdings

As shown before, chequable deposits undergo multiple expansion, whereas currency does not. Hence, when chequable deposits are converted into currency, as long as the amount of excess reserves is held constant, a switch is made from a component of the money supply that undergoes multiple expansion to one that does not. The overall level of multiple expansion declines, and the money supply falls. However, if currency holdings fall, a switch is made into chequable deposits that undergo multiple deposit expansion, so the money supply rises. This analysis suggests the following result: ***Holding excess reserves constant, the money supply is negatively related to currency holdings.***

15.6 OVERVIEW OF THE MONEY SUPPLY PROCESS

LO 15.6 Summarize how the "three players" can influence the money supply.

We now have a model of the money supply process in which all three of the players—the Bank of Canada, depositors, and banks—directly influence the money supply. As a study aid, Summary Table 15-1 charts the money supply responses to the five factors discussed above and gives a brief synopsis of the reasoning behind them.

The variables are grouped by the player who is the primary influence behind the variable. The Bank of Canada, for example, influences the money supply by controlling the first variable. Depositors influence the money supply through their decisions about holdings of currency, while banks influence the money supply with their decisions about borrowings from the Bank of Canada and excess reserves.

SUMMARY TABLE 15-1

Money Supply Response

Player	Variable	Change in Variable	Money Supply Response	Reason
Bank of Canada	Nonborrowed monetary base, MB_n	↑	↑	More *MB* for deposit creation
Banks	Desired reserve ratio, r_d	↑	↓	Less multiple deposit expansion
	Borrowed reserves, *BR*	↑	↑	More *MB* for deposit creation
	Excess reserves	↑	↓	Less loans and deposit creation
Depositors	Currency holdings	↑	↓	Less multiple deposit expansion

Note: Only increases (↑) in the variables are shown. The effects of decreases on the money supply would be the opposite of those indicated in the "Money Supply Response" column.

15.7 THE MONEY MULTIPLIER

LO 15.7 Calculate and interpret changes in the money multiplier.

The intuition inherent in the preceding section is sufficient for you to understand how the money supply process works. For those of you who are more mathematically inclined, we can derive all of the above results using a concept called the **money multiplier**, denoted by *m*, which tells us how much the money supply changes for a given change in the monetary base. The relationship between the money supply *M*, the money multiplier, and the monetary base is described by the following equation:

$$M = m \times MB \tag{2}$$

The money multiplier *m* tells us what multiple of the monetary base is transformed into the money supply. Because the money multiplier is typically larger than one, the alternative name for the monetary base, *high-powered money*, is logical: A $1 change in the monetary base typically leads to more than a $1 change in the money supply.

Deriving the Money Multiplier

Let's assume that the desired holdings of currency *C* and excess reserves *ER* grow proportionally with chequable deposits *D*; in other words, we assume that the ratios of these items to chequable deposits are constants in equilibrium, as the braces in the following expressions indicate:

$$c = \{C/D\} = \text{currency ratio}$$
$$e = \{ER/D\} = \text{excess reserves ratio}$$

We will now derive a formula that describes how the currency ratio desired by depositors, the excess reserves ratio desired by banks, and the desired reserve ratio set by the banks affect the multiplier *m*. We begin the derivation of the model of the money supply with the following equation:

$$R = DR + ER$$

which states that the total amount of reserves in the banking system *R* equals the sum of desired reserves *DR* and excess reserves *ER*. (Note that this equation corresponds to the equilibrium condition $DR = R$ given earlier in the chapter, where excess reserves were assumed to be zero.)

The total amount of desired reserves equals the desired reserve ratio r_d times the amount of chequable deposits *D*:

$$DR = r_d \times D$$

Substituting $r_d \times D$ for *DR* in the first equation yields an equation that links reserves in the banking system to the amount of chequable deposits and excess reserves they can support:

$$R = (r_d \times D) + ER$$

A key point here is that banks set the desired reserve ratio r_d to less than one. Thus \$1 of reserves can support more than \$1 of deposits, and the multiple expansion of deposits can occur.

Let's see how this works in practice. If excess reserves are held at zero $(ER = 0)$, the desired reserve ratio is set at $r_d = 0.10$, and the level of chequable deposits in the banking system is \$1600 billion, then the amount of reserves needed to support these deposits is \$160 billion $(= 0.10 \times \$1600 \text{ billion})$. The \$160 billion of reserves can support 10 times this amount in chequable deposits because multiple deposit creation will occur.

Because the monetary base *MB* equals currency *C* plus reserves *R*, we can generate an equation that links the amount of the monetary base to the levels of chequable deposits and currency by adding currency to both sides of the preceding equation:

$$MB = R + C = (r_d \times D) + ER + C$$

Notice that this equation reveals the amount of the monetary base needed to support the existing amounts of chequable deposits, currency, and excess reserves.

To derive the money multiplier formula in terms of the currency ratio $c = \{C/D\}$ and the excess reserves ratio $e = \{\text{ER}/D\}$, we rewrite the last equation, specifying *C* as $c \times D$ and *ER* as $e \times D$:

$$MB = (r_d \times D) + (e \times D) + (c \times D) = (r_d + e + c) \times D$$

We next divide both sides of the equation by the term inside the parentheses to get an expression linking chequable deposits *D* to the monetary base *MB*:

$$D = \frac{1}{r_d + e + c} \times MB \tag{3}$$

Using the M1+ definition of the money supply as currency plus chequable deposits $(M = D + C)$ and again specifying *C* as $c \times D$, we get

$$M = D + (c \times D) = (1 + c) \times D$$

Substituting in this equation the expression for D from Equation 3 yields

$$M = \frac{1 + c}{r_d + e + c} \times MB \tag{4}$$

We have derived an expression in the form of our earlier Equation 2. As you can see, the ratio that multiplies MB is the money multiplier, which tells how much the money supply changes in response to a given change in the monetary base (high-powered money). The money multiplier m is thus

$$m = \frac{1 + c}{r_d + e + c} \tag{5}$$

It is a function of the currency ratio set by depositors c, the excess reserves ratio set by banks e, and the desired reserve ratio r_d, also set by banks.

Intuition Behind the Money Multiplier

To get a feel for what the money multiplier means, let's construct a numerical example with realistic numbers for the following variables:

r_d = desired reserve ratio = 0.10
C = currency in circulation = \$1200 billion
D = chequable deposits = \$1600 billion
ER = excess reserves = \$2500 billion
M = money supply (M1+) = $C + D$ = \$2800 billion

From these numbers we can calculate the values for the currency ratio c and the excess reserves ratio e:

$$c = \frac{\$1200 \text{ billion}}{\$1600 \text{ billion}} = 0.75$$
$$e = \frac{\$2500 \text{ billion}}{\$1600 \text{ billion}} = 1.56$$

The resulting value of the money multiplier is

$$m = \frac{1 + 0.75}{0.1 + 1.56 + 0.75} = \frac{1.75}{2.41} = 0.73$$

The money multiplier of 0.73 tells us that, given the desired reserve ratio of 10% on chequable deposits and the behaviour of depositors, as represented by $c = 0.75$, and banks, as represented by $e = 1.56$, a \$1 increase in the monetary base leads to a \$0.73 increase in the money supply (M1+).

An important characteristic of the money multiplier is that it is far less than the simple deposit multiplier of 10 found earlier in the chapter. There are two reasons for this result. First, ***although deposits undergo multiple expansion, currency does not.*** Thus, if some portion of the increase in high-powered money finds its way into currency, this portion does not undergo multiple deposit expansion. In our simple model earlier in the chapter, we did not allow for this possibility, and so the increase in reserves led to the maximum amount of multiple deposit creation. However, in our current model of the money multiplier, the

level of currency does rise when the monetary base *MB* and chequable deposits *D* increase, because *c* is greater than zero. As previously stated, any increase in *MB* that goes into an increase in currency is not multiplied, so only part of the increase in *MB* is available to support chequable deposits that undergo multiple expansion. The overall level of multiple deposit expansion must be lower, meaning that the increase in *M*, given an increase in *MB*, is smaller than indicated by the simple model earlier in the chapter.

Second, since *e* is positive, ***any increase in the monetary base and deposits leads to higher excess reserves.*** When there is an increase in *MB* and *D*, the resulting increase in excess reserves means that the amount of reserves used to support chequable deposits does not increase as much as it otherwise would. Hence the increase in chequable deposits and the money supply are lower, and the money multiplier is smaller.

Prior to 2008, in the United States, the excess reserves ratio *e* was almost always very close to zero (less than 0.001), and so its impact on the money multiplier (Equation 5) was essentially irrelevant. When *e* is close to zero, the money multiplier is always greater than one, and it was around 1.6 during that period. However, as we will see in the next chapter, nonconventional monetary policy during the global financial crisis, and again during the coronavirus pandemic, caused excess reserves to skyrocket to over $2 trillion. Such an extraordinarily large value of *e* caused the excess reserves factor in the money multiplier equation to become dominant, and so the U.S. money multiplier fell to below one, as discussed above.

Money Supply Response to Changes in the Factors

By recognizing that the monetary base is $MB = MB_n + BR$, we can rewrite Equation 2 as

$$M = m \times (MB_n + BR) \tag{6}$$

Now we can demonstrate algebraically all the results given in Summary Table 15-1, which shows how the money supply responds to the changes in the factors.

As you can see from Equation 6, a rise in MB_n or *BR* raises the money supply *M* because the money multiplier *m* is always greater than zero. We can see that a rise in the desired reserve ratio r_d lowers the money supply by calculating the value of the money multiplier (using Equation 5) in our numerical example when r_d increases from 10% to 15% (leaving all other variables unchanged). The money multiplier then becomes

$$m = \frac{1 + 0.75}{0.15 + 1.56 + 0.75} = \frac{1.75}{2.46} = 0.71$$

which, as we would expect, is less than 0.73.

Similarly, we can see that a rise in excess reserves lowers the money supply by calculating the money multiplier when *e* is increased from 1.56 to 3.0. The money multiplier decreases from 0.73 to

$$m = \frac{1 + 0.75}{0.1 + 3.00 + 0.75} = \frac{1.75}{3.85} = 0.45$$

We can also analyze what happens in our numerical example when there is a rise in the currency ratio *c* from 0.75 to 1.50. In this case, something peculiar happens. Instead of falling, the money multiplier rises from 0.73 to

$$m = \frac{1 + 1.50}{0.1 + 1.56 + 1.50} = \frac{2.50}{3.20} = 0.78$$

At first glance, this result might seem counterintuitive. After all, a dollar of monetary base that goes into currency only increases the money supply by one dollar,

whereas a dollar of monetary base that goes into deposits leads to multiple deposit expansion that increases the money supply by a factor of 10. Thus it seems as though the shift from deposits to currency should lower the overall amount of multiple expansion and hence the money supply. This reasoning is correct, but it assumes a small value of the excess reserves ratio. Indeed, that is the case during normal times, when the excess reserves ratio is near zero. However, in our current situation, in which the excess reserves ratio e is abnormally high, when a dollar moves from deposits into currency, the amount of excess reserves falls by a large amount, which releases reserves to support more deposits, causing the money multiplier to rise.[7]

APPLICATION Quantitative Easing and the Money Supply During the Global Financial and the Coronavirus Crises

When the global financial crisis became virulent in September 2008, the Federal Reserve in the United States initiated lending programs and large-scale asset-purchase programs in an attempt to bolster the U.S. economy. By the end of 2014, these lending programs and purchases of securities had led to a 350% increase in the Fed's balance sheet and the monetary base. Because these lending and asset-purchase programs, discussed further in Chapter 16, resulted in a huge expansion of the monetary base, they have been given the name "quantitative easing." As our analysis in this chapter indicates, such a massive expansion of the monetary base could potentially lead to a large expansion of the money supply. However, as shown in Figure 15-1, when the monetary base increased by 350%, the M1 money supply rose by only 100%. How can we explain this result using our money supply model?[8]

The answer is that despite the huge increase in the monetary base, the money supply rose by much less because the money multiplier fell by over 50%. To explain this decline in the money multiplier, let's look at Figure 15-2, which shows the currency ratio c and the excess reserves ratio e in the United States for 2007–2020. We see that the currency ratio had a slight downward trend, which would have raised, not lowered, the money multiplier. Instead, we have to look to the extraordinary rise in the excess reserves ratio e, which climbed 30-fold from September 2008 to December 2014.

What explains this substantial increase in the excess reserves ratio e from September 2008 to December 2014? When the Federal Reserve began to pay interest on excess reserves in October 2008, the interest rate on these reserves was either equal to or slightly greater than the rate at which the banks could lend them out in the federal

[7]All the above results can be derived more generally from the Equation 5 formula for m as follows. When r_d or e increases, the denominator of the money multiplier increases, and therefore the money multiplier must decrease. As long as $r_d + e$ is less than one (as is usually the case), an increase in c raises the denominator of the money multiplier proportionally by more than it raises the numerator. The increase in c causes the money multiplier to fall. However, when $r_d + e$ is greater than one (the current situation), an increase in c raises the numerator of the money multiplier proportionally by more than it raises the denominator, so the money multiplier rises. Recall that the money multiplier in Equation 5 is for the M1+ definition of money. The second online appendix to Chapter 15 discusses how the multiplier for M2+ is determined.

[8]If you would like to see a similar application of the money supply model to what happened to the money supply during the Great Depression period from 1930 to 1933, go to the third online appendix to Chapter 15.

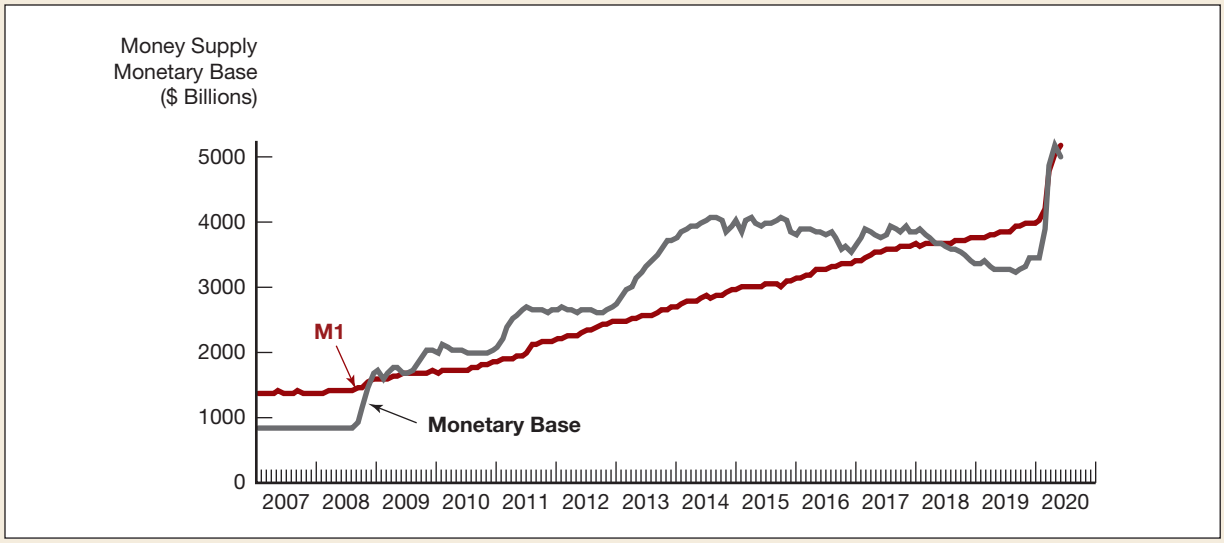


FIGURE 15-1 **M1 and the Monetary Base in the United States, 2007–2020**

In percentage terms the money supply rose by substantially less than the monetary base during the two quantitative easing episodes, the global financial crisis and the coronavirus pandemic.

Source: Federal Reserve Bank of St. Louis FRED database: fred.stlouisfed.org/series/BOGMBASE; https://fred.stlouisfed.org/series/M1SL.

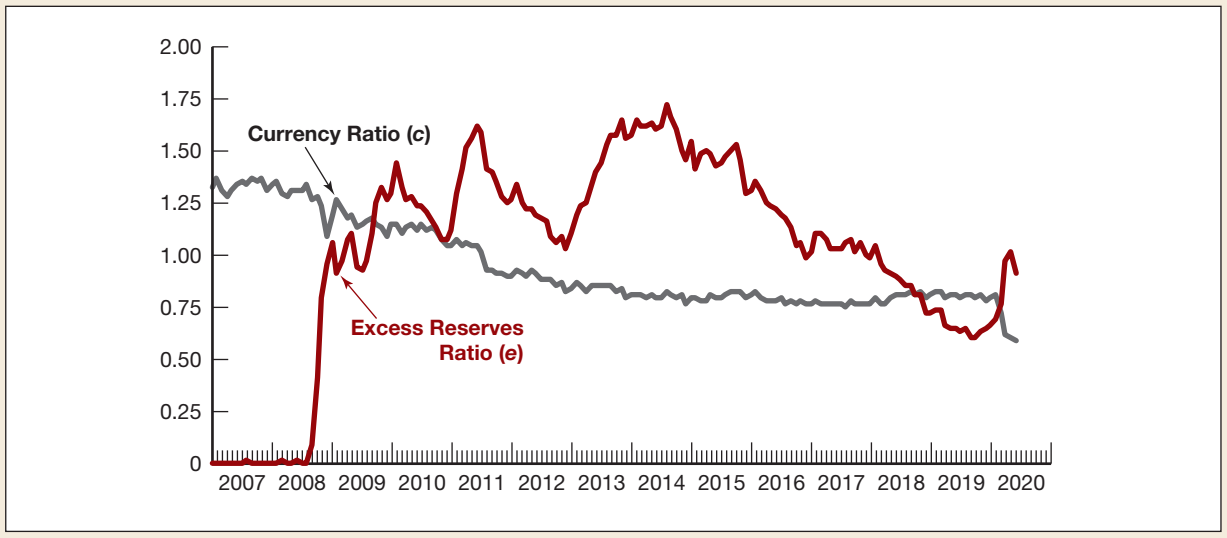


FIGURE 15-2 **Excess Reserves and Currency Ratio in the United States, 2007–2020**

The currency ratio c was relatively steady during this period, whereas the excess reserves ratio e rose sharply after quantitative easing during the global financial and coronavirus crises.

Source: Federal Reserve Bank of St. Louis FRED database: fred.stlouisfed.org/series/EXCSRESNS; https://fred.stlouisfed.org/series/CURRCIR; https://fred.stlouisfed.org/series/TCDNS.

funds market. Because excess reserves now had an opportunity cost close to zero, when the Fed's actions created far more reserves than were needed for banks to meet their reserve requirements, banks were willing to hold any amount of these excess reserves. Therefore, the increase in reserves from quantitative easing led to a large increase in the excess reserves ratio *e*. As predicted by our money supply model, the huge increase in *e* sharply lowered the money multiplier, and so the money supply did not undergo that large an expansion from August 2008 to December 2014, despite the huge increase in the monetary base.

When the coronavirus pandemic struck in March 2020, the Federal Reserve again engaged in massive quantitative easing programs to prevent a collapse of the economy. The effects on the monetary base and the money supply displayed a similar pattern to that which occurred when the global financial crisis struck. From February 2020 to June 2020, the monetary base increased by 45%, while the money supply only increased by a much smaller 29%. Just as in the aftermath of the global financial crisis, the increase in excess reserves led to the excess reserves ratio rising sharply: From February 2020 to June 2020, the excess reserves ratio climbed by 33%. As predicted by our money supply model, this increase in *e* led to a decline in the money multiplier, so the increase in the monetary base again did not lead to as large an increase in the money supply. ◆

SUMMARY

1. The three players in the money supply process are the central bank, banks (depository institutions), and depositors.
2. Four items in the Bank of Canada's balance sheet are essential to our understanding of the money supply process: the two liability items, currency in circulation and reserves, which together make up the monetary base; and the two asset items, securities and loans to financial institutions.
3. The Bank of Canada controls the monetary base through open market operations and extensions of loans to financial institutions and has better control over the monetary base than over reserves. Although float and government deposits with the Bank of Canada undergo substantial short-run fluctuations, which complicate control of the monetary base, they do not prevent the Bank of Canada from accurately controlling it.
4. A single bank can make loans up to the amount of its excess reserves, thereby creating an equal amount of deposits. The banking system can create a multiple expansion of deposits, because as each bank makes a loan and creates deposits, the reserves find their way to another bank, which uses them to make loans and create additional deposits. In the simple model of multiple deposit creation, in which banks do not hold on to excess reserves and the public holds no currency, the multiple increase in chequable deposits (simple deposit multiplier) equals the reciprocal of the desired reserve ratio.
5. The simple model of multiple deposit creation has serious deficiencies. Decisions by depositors to increase their holdings of currency or by banks to hold excess reserves will result in a smaller expansion of deposits than is predicted by the simple model. All three players—the Bank of Canada, banks, and depositors—are important in the determination of the money supply.
6. The money supply is positively related to the nonborrowed monetary base MB_n, which is determined by open market operations, and the level of borrowed reserves (lending) from the Bank of Canada, *BR*. The money supply is negatively related to the desired reserve ratio, r_d, and excess reserves. The money supply is also negatively related to holdings of currency but only if excess reserves do not vary much when there is a shift between deposits and currency. The model of the money supply process takes into account the behaviour of all three players in the money supply process: the Bank of Canada through open market

operations, banks through their decisions to borrow from the Bank of Canada and hold excess reserves, and depositors through their decisions about the holding of currency.

7. The monetary base is linked to the money supply using the concept of the money multiplier, which tells us how much the money supply changes when the monetary base changes.

MyLab Economics Make the grade with MyLab Economics: Select end-of-chapter questions can be found on MyLab Economics. You can practise them as often as you want, and most feature step-by-step guided instructions to help you find the right answer.

QUESTIONS

Unless otherwise noted, the following assumptions are made in all questions: The desired reserve ratio on chequable deposits is 10%, banks do not hold any excess reserves, and the public's holdings of currency do not change.

1. Classify each of these transactions as an asset, a liability, or neither for each of the "players" in the money supply process—the Bank of Canada, banks, and depositors.
   a. You get a $10 000 loan from the bank to buy an automobile.
   b. You deposit $400 into your chequing account at the local bank.
   c. The Bank of Canada provides an emergency loan to a bank for $1 000 000.
   d. A bank borrows $500 000 in overnight loans from another bank.
   e. You use your debit card to purchase a meal at a restaurant for $100.
2. The First National Bank receives an extra $100 of reserves but decides not to lend out any of these reserves. How much deposit creation takes place for the entire banking system?
3. Suppose the Bank of Canada buys $1 million of bonds from the First National Bank. If the First National Bank and all other banks use the resulting increase in reserves to purchase securities only and not to make loans, what will happen to chequable deposits?
4. If a bank depositor withdraws $1000 of currency from an account, what happens to reserves, chequable deposits, and the monetary base?
5. If a bank sells $10 million of bonds to the Bank of Canada to pay back $10 million on the loan it owes, what is the effect on the level of chequable deposits?
6. If you decide to hold $100 less cash than usual and therefore deposit $100 more cash in the bank, what effect will this have on chequable deposits in the banking system if the rest of the public keeps its holdings of currency constant?
7. "The Bank of Canada can perfectly control the amount of reserves in the system." Is this statement true, false, or uncertain? Explain.
8. "The Bank of Canada can perfectly control the amount of the monetary base but has less control over the composition of the monetary base." Is this statement true, false, or uncertain? Explain.
9. The Bank of Canada buys $100 million of bonds from the public, and banks also lower their desired reserve ratio. What will happen to the money supply?
10. Describe how each of the following can affect the money supply: (a) the central bank, (b) banks, and (c) depositors.
11. "The money multiplier is necessarily greater than one." Is this statement true, false, or uncertain? Explain your answer.
12. What effect might a financial panic have on the money multiplier and the money supply? Why?
13. During the Great Depression years from 1930 to 1933, both the currency ratio c and the excess reserves ratio e rose dramatically. What effect did these factors have on the money multiplier?

14. In October 2008, the Federal Reserve began paying interest on the amount of excess reserves held by banks. How, if at all, might this affect the multiplier process and the money supply?

15. The money multiplier declined significantly during the period 1930–1933 and also during the recent financial crisis of 2008–2010. Yet the M1 money supply *decreased* by 25% in the Depression period but *increased* by more than 20% during the recent financial crisis. What explains the difference in outcomes?

APPLIED PROBLEMS

Unless otherwise noted, the following assumptions are made in all of the applied problems: The desired reserve ratio on chequable deposits is 10%, banks do not hold any excess reserves, and the public's holdings of currency do not change.

16. If the Bank of Canada sells $2 million of bonds to the First National Bank, what happens to reserves and the monetary base? Use T-accounts to explain your answer.

17. If the Bank of Canada sells $2 million of bonds to Irving the Investor, who pays for the bonds with a briefcase filled with currency, what happens to reserves and the monetary base? Use T-accounts to explain your answer.

18. If the Bank of Canada lends five banks a total of $100 million, but depositors withdraw $50 million and hold it as currency, what happens to reserves and the monetary base? Use T-accounts to explain your answer.

19. Using T-accounts, show what happens to chequable deposits in the banking system when the Bank of Canada lends $1 million to the First National Bank.

20. Using T-accounts, show what happens to chequable deposits in the banking system when the Bank of Canada sells $2 million of bonds to the First National Bank.

21. If the Bank of Canada buys $1 million of bonds from the First National Bank, but an additional 10% of any deposit is held as excess reserves, what is the total increase in chequable deposits? (*Hint:* Use T-accounts to show what happens at each step of the multiple expansion process.)

22. If reserves in the banking system increase by $1 billion because the Bank of Canada lends $1 billion to financial institutions, and chequable deposits increase by $9 billion, why isn't the banking system in equilibrium? What will continue to happen in the banking system until equilibrium is reached? Show the T-account for the banking system in equilibrium.

23. If the Bank of Canada reduces reserves by selling $5 million worth of bonds to the banks, what will the T-account of the banking system look like when the banking system is in equilibrium? What will have happened to the level of chequable deposits?

24. If the Bank of Canada sells $1 million of bonds and banks reduce their borrowings from the Bank of Canada by $1 million, predict what will happen to the money supply.

25. Suppose that currency in circulation is $600 billion, the amount of chequable deposits is $900 billion, and excess reserves are $15 billion.

a. Calculate the money supply, the currency deposit ratio, the excess reserve ratio, and the money multiplier.

b. Suppose the central bank conducts an unusually large open market purchase of bonds held by banks of $1400 billion due to a sharp contraction in the economy. Assuming the ratios you calculated in part (a) remain the same, predict the effect on the money supply.

c. Suppose the central bank conducts the same open market purchase as in part (b) except that banks choose to hold all of these proceeds as excess reserves rather than loan them out due to fear of a financial crisis. Assuming that currency and deposits remain the same, what happens to the amount of excess reserves, the excess reserve ratio, the money supply, and the money multiplier?

d. During the financial crisis in 2008, the Federal Reserve began injecting the banking system with massive amounts of liquidity, and at the same time, very little lending occurred. As a result, the M1 money multiplier was below one for most of the time from October 2008 through 2011. How does this scenario relate to your answer to part (c)?

DATA ANALYSIS PROBLEMS

1. Go to the CANSIM webpage and find monthly data (from January 1968 to December 2020) on the monetary base (series V37145), M1++ gross (series V37152), and M2++ gross (series V41552801).
    a. What is the average M1++ growth rate? Its variance?
    b. What is the average M2++ growth rate? Its variance?
    c. Which monetary aggregate is more variable?
    d. Calculate the M1++ money multiplier. What is its average value?
    e. Calculate the M2++ money multiplier. What is its average value?
    f. Compare the two money multipliers.
    g. Does it appear that the money multipliers are increasing or decreasing?
2. **Real-Time Data Analysis** Go to the St. Louis Federal Reserve FRED database and find the most current data available on Currency (CURRNS), Total Checkable Deposits (TCDNS), and Excess Reserves (EXCSRESNS).
    a. Calculate the value of the currency deposit ratio c.
    b. Calculate the value of the excess reserve ratio, e.
    c. Assuming a desired reserve ratio r_d of 11%, calculate the value of the money multiplier m.
3. **Real-Time Data Analysis** Go to the St. Louis Federal Reserve FRED database and find data on the M1 Money Stock (M1SL) and the Monetary Base (BOGMBASE).
    a. Calculate the value of the money multiplier using the most recent data available and the data from five years prior.
    b. Based on your answer to part (a), how much would a $100 million open market purchase of securities affect the M1 money supply today and five years ago?

16 Tools of Monetary Policy

Learning Objectives

16.1 Characterize the framework for the implementation of monetary policy in Canada (the LVTS, the target and the operating band for the overnight interest rate, and the Bank of Canada's standing liquidity facilities).

16.2 Illustrate the market for reserves, and demonstrate how changes in monetary policy can affect the equilibrium overnight interest rate.

16.3 Discuss the Bank of Canada's approach to monetary policy.

16.4 Summarize how conventional monetary policy tools are implemented and the relative advantages and limitations of each tool.

16.5 Explain the key monetary policy tools that are used when conventional policy is no longer effective.

16.6 Identify the distinctions and similarities between the monetary policy tools of the Bank of Canada and those of the Federal Reserve and the European Central Bank.

Preview

In this chapter, we examine the tools of monetary policy used by the Bank of Canada to control the money supply and interest rates. Because the Bank's use of these policy tools has such an important impact on interest rates and economic activity, it is vital to understand how the Bank of Canada wields them in practice and how relatively useful each tool is.

In recent years, the Bank of Canada has increasingly focused on the **overnight rate** (the interest rate on overnight loans of reserves from one bank to another) as the primary instrument of monetary policy. Since December 2000, the Bank of Canada has announced an overnight rate target eight times throughout the year, an announcement that is watched closely by market participants because it affects interest rates throughout the economy. Thus, to fully understand how the Bank of Canada uses its tools in the conduct of monetary policy, we must understand not only their effects on the money supply but also their direct effect on the overnight interest rate and how they can be used to help achieve an overnight interest rate that is close to the target. The chapter therefore begins with the institutional framework within which the Bank of Canada conducts monetary policy, followed by a supply and demand analysis of the market for reserves to explain how the Bank of Canada's setting for the three monetary policy tools—open market operations, settlement balances management, and Bank of Canada lending—determine the overnight interest rate. We next go on to look in more detail at each of these tools of monetary policy to see how they are used in practice and to gain a sense of their relative advantages. We then examine the nonconventional tools of monetary policy that the Bank of Canada and other central banks were forced to use during the extraordinary circumstances of the 2007–2009 financial crisis and the coronavirus pandemic in 2020. The chapter ends with a discussion of the tools of monetary policy used by other central banks besides the Bank of Canada.

16.1 THE FRAMEWORK FOR THE IMPLEMENTATION OF MONETARY POLICY

LO 16.1 Characterize the framework for the implementation of monetary policy in Canada (the LVTS, the target and the operating band for the overnight interest rate, and the Bank of Canada's standing liquidity facilities).

The tools used by the Bank of Canada to implement monetary policy are closely linked to the institutional arrangements regarding the clearing and settlement systems in the

Canadian economy. Understanding the tools of monetary policy therefore requires that we know the key features of the framework for the implementation of monetary policy. As you will see, this framework has been designed to encourage deposit-taking financial institutions to deal directly with the market rather than with the Bank of Canada.[1]

The Large Value Transfer System (LVTS)

The core of the Canadian payments system is the **Large Value Transfer System (LVTS)**, introduced on February 4, 1999, and operated by Payments Canada. In 2019, Payments Canada cleared and settled a daily average value of about $218 billion and a total value of over $55 trillion, with the LVTS accounting for a daily average transaction value of about $190 billion and a total value of close to $48 trillion. The LVTS is an electronic, real-time net settlement network designed to provide immediate finality and settlement to time-critical transactions. In addition to the Bank of Canada, currently there are 15 direct **LVTS participants**.[2] A financial institution can participate directly in the LVTS if it is a member of Payments Canada, uses the SWIFT telecommunications network, maintains a settlement account at the Bank of Canada, and has agreements regarding borrowing from the Bank of Canada and pledging eligible collateral. All other members of Payments Canada that are not direct participants in the LVTS are able to arrange LVTS payments through the LVTS participants.

The LVTS has been put in place in order to eliminate **systemic risk**—the risk to the entire payments system due to the inability of one financial institution to fulfill its payment obligations in a timely fashion. In fact, each LVTS payment is subject to real-time risk-control tests to confirm that sufficient collateral is available and is final and irrevocable in real time. In particular, LVTS participants can make a payment only if they have, in real time (right now), either positive settlement balances in their accounts with the Bank of Canada, posted collateral (such as Government of Canada Treasury bills and bonds), or explicit lines of credit with other participants. As a result, the system will settle at the end of each day even in the face of risk and liquidity problems.

Of course, it is not just Canada that is concerned about systemic risk. Real-time settlement systems were implemented by Sweden in 1986, Germany and Switzerland in 1987, Japan in 1988, Italy in 1989, Belgium and the United Kingdom in 1996, and by France, Hong Kong, and the Netherlands in 1997. Moreover, the central banks of the G-10 countries, through the Bank for International Settlements, have developed minimum standards for the operation of the global payment network for large-value funds transfers.

In Canada's LVTS, participants know in real time their large-value, wholesale transactions (over $50 000). Although these transactions account for less than 1% of the total number of transactions, they account for about 87% of the value of transactions. For example, the LVTS is used to settle close to 40 000 payments daily, with a value close to $190 billion. This information eliminates most of the uncertainty from settlement balance prediction—the largest reason for financial institutions not being able to hit their target settlement balances with the Bank of Canada in the pre-LVTS system.

[1]For more details, see Donna Howard, "A Primer on the Implementation of Monetary Policy in the LVTS Environment," *Bank of Canada Review* (Autumn, 1998): 57–66; Kevin Clinton, "Implementation of Monetary Policy in a Regime with Zero Reserve Requirements," Bank of Canada Working Paper: 97–98; and "The Framework for the Implementation of Monetary Policy in the Large Value Transfer System Environment," Bank of Canada Release, March 31, 1999.

[2]These are the Big Six (Bank of Montreal, Bank of Nova Scotia, Canadian Imperial Bank of Commerce, National Bank of Canada, Royal Bank of Canada, and Toronto-Dominion Bank), ATB Financial, Bank of America National Association, BNP Paribas, Federation des caisses Desjardins du Québec, Central 1 Credit Union, HSBC Bank Canada, ICICI Bank Canada, Laurentian Bank of Canada, and State Street.

Non-LVTS (ACSS) Transactions

Although the LVTS eliminates the uncertainty from daily wholesale settlement balances prediction, there is still a residual stochastic element in settlement balances from non-LVTS paper-based payments items (such as paper cheques, travellers' cheques, gift certificates, and money orders) and small-value electronic payments items (such as pre-authorized debits and fund transfers, direct deposit items, and debit card transactions). Those items are cleared through the Automated Clearing Settlement System (ACSS), an electronic payments system also owned and operated by Payments Canada but without the vigorous risk controls of the LVTS. The ACSS aggregates interbank payments and calculates the net amounts to be transferred from and to each participant's settlement account with the Bank of Canada. The Bank completes the settlement the next day (at midday) through the LVTS. In 2019, the ACSS supported 99% of the daily average transaction volume (about 33 million items) and 13% of the daily average transaction value (about $29 billion) cleared by all Payments Canada systems.

Regarding the smaller value of ACSS settlements, **direct clearers** with positive clearing balances in the ACSS (that is, with net credit positions) receive an interest payment (calculated daily by Payments Canada at the Bank of Canada's target overnight interest rate) to cover their cost of crediting their customers' accounts on the business day before receiving funds in settlement. On the other hand, direct clearers with net debit positions in the ACSS make payments (also at the Bank's target overnight interest rate) to offset their benefit of debiting the accounts of their customers on the day before paying funds in settlement. Interest compensation is included in the clearing balances settled at the Bank of Canada using LVTS funds.[3]

The Bank of Canada's Policy Rate

The settlement of payment obligations among LVTS participants provides the setting for the conduct of monetary policy by the Bank of Canada. In particular, throughout any given business day, LVTS participants send payments to, and receive payments from, each other. The LVTS uses **multilateral netting**, in which only the net credit or debit position of each participant vis-à-vis all other participants is calculated for settlement, thereby reducing the need for a large amount of settlement balances. Thus, at the end of the day, some LVTS members may be short of settlement balances while others may have settlement balances left over. To settle these differences, they borrow from and lend to each other, on a one-day (or overnight) basis, in the overnight money market.

This overnight money market is the key market for finance and monetary policy in Canada. It is very liquid, with a broad range of participants, the most active of which are deposit-taking institutions and their investment dealer affiliates.[4] The interest rate at which participants borrow and lend overnight funds to each other in the money market is called the **overnight interest rate**, i_{or}. This rate is the shortest-term rate available and forms the base of any term structure of interest rates relation.

The Bank of Canada signals its monetary policy stance by announcing a target for the overnight interest rate (the overnight interest rate is known as the **reference rate**). The target for the overnight rate, known as the **policy rate**, is the main tool the

[3]There are 11 ACSS direct clearers in addition to the Bank of Canada—the Big Six, ATB Financial, Federation des caisses Desjardins du Québec, Central 1 Credit Union, HSBC Bank Canada, and the Laurentian Bank of Canada. Direct clearers maintain a settlement account at the Bank of Canada and participate directly in the ACSS.

[4]See Eugene Lundrigan and Sari Toll, "The Overnight Market in Canada," *Bank of Canada Review* (Winter 1997–1998): 27–42, for details regarding the evolution of the overnight market.

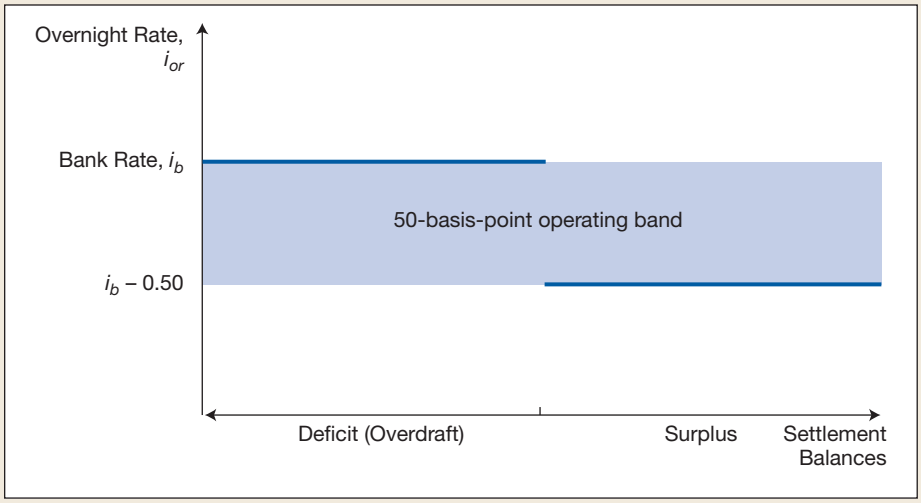


FIGURE 16-1 Operating Band for the Overnight Interest Rate

Bank of Canada uses to conduct monetary policy. This rate refers to collateralized, market-based overnight transactions. The implementation of the stance of monetary policy is more successful when the reference rate is closer to the policy rate—that is, the closer the overnight rate is to the target for the overnight interest rate.

The Operating Band for the Overnight Interest Rate

The Bank of Canada implements monetary policy by changing the policy rate in order to influence other short-term interest rates and the exchange rate. In fact, in normal times the Bank's operational objective is to keep the overnight rate within an **operating band** (also known as a *channel* or *corridor*) of 50 basis points (½ of 1%). As Figure 16-1 shows, the upper limit of the operating band defines the bank rate, i_b. ***The bank rate is the interest rate the Bank of Canada charges LVTS participants that require an overdraft loan (advance) to cover negative settlement balances on the books of the Bank at the end of the banking day***. The lower limit of the operating band is the rate the Bank pays to LVTS participants with positive settlement balances at the end of the day. The midpoint of the operating band is the Bank of Canada's target for the overnight rate, the operating target of the Bank of Canada's monetary policy. When, for example, the operating band is from 1.5% to 2.0%, the bank rate is 2.0%, the rate the Bank pays on deposits to LVTS participants is 1.5%, and the Bank's target (which is the policy rate referred to earlier) for the overnight interest rate is 1.75%.

Since December 2000, the Bank of Canada has operated under a system of eight "fixed" dates throughout the year for announcing any changes to the target and the operating band for the overnight rate and has reserved the option of acting between the fixed dates in "extraordinary circumstances." Early in the morning (at 9:00 a.m.) on each of the eight announcement dates, the Bank of Canada announces the target and an operating band of 50 basis points for the overnight interest rate. (See the FYI box "Monetary Policy Implementation in the LVTS Environment.")

FYI Monetary Policy Implementation in the LVTS Environment

In the LVTS environment, the Bank of Canada operates under a system of eight fixed dates throughout the year for announcing, via a press release (at 9:00 a.m.), the target and the operating band for the overnight interest rate. The upper limit of the operating band defines the bank rate, i_b, at which the Bank provides overdraft loans to LVTS participants at the end of the day, and the lower limit, $i_b - 0.50$, is the rate the Bank pays to LVTS participants with positive settlement balances at the end of the day.

The operating band for the overnight rate, reinforced by the Bank's standing liquidity facilities, and a target level of settlement balances (as of March 19, 2020, the target level of settlement balances is $2 billion) are currently the framework within which the Bank of Canada implements monetary policy. The midpoint of the operating band, the policy rate, is the operating target of the Bank's monetary policy.

In targeting the midpoint of the operating band, the Bank of Canada will intervene in the overnight market at 11:45 a.m. to conduct open market buyback operations at the target overnight interest rate, if required. In particular, the Bank will use SPRAs if overnight funds are traded above the target for the overnight rate and SRAs if overnight funds are traded below the target for the overnight rate. There will be no Bank of Canada intervention in the overnight market if the overnight rate is around the target rate.

To maintain the desired level of settlement balances, the Bank of Canada must neutralize the net impact of its open market buyback operations and of certain federal government and Bank of Canada flows that potentially affect settlement balances. The neutralization is effected through the shifting (transfer) of federal government deposits (balances) between the government's account at the Bank of Canada and the participants in the auction of government cash balances. The shifting is made through the twice-daily auctions of Receiver General (federal government) balances (the first at 9:15 a.m. and the second at 4:15 p.m.).

The LVTS has a pre-settlement trading period of half an hour at the end of the banking day (6:00–6:30 p.m.) to permit participating financial institutions to adjust positions with each other (in the overnight market) at a better return than can be achieved at the Bank of Canada's standing facilities. LVTS participants with settlement imbalances at the end of the banking day use the Bank of Canada's standing facilities to bring their settlement balances to the target level. That is, LVTS participants that require an overdraft loan to cover negative settlement balances on the books of the Bank borrow from the Bank at the bank rate, i_b. LVTS participants with positive settlement balances at the end of the day earn (during normal times) the bank rate less 50 basis points, $i_b - 0.50$.

The Market Timetable

9:00 a.m.	Bank of Canada announces changes (if any) to the target and the operating band for the overnight rate
9:15 a.m.	Auction for (federal) government cash balances
11:45 a.m.	Open market buyback operations (if required)
4:15 p.m.	Auction for (federal) government cash balances
6:00 p.m.	Close of LVTS for client transactions
6:00–6:30 p.m.	Pre-settlement period
8:00 p.m. or earlier	Settlement of LVTS balances at the Bank of Canada

Sources: Donna Howard, "A Primer on the Implementation of Monetary Policy in the LVTS Environment," *Bank of Canada Review* (Autumn 1998): 57–66; and "The Framework for the Implementation of Monetary Policy in the Large Value Transfer System Environment," *Bank of Canada Release*, March 31, 1999.

A similar channel/corridor system for setting the overnight interest rate is also used by other central banks, including the Federal Reserve, the European Central Bank, the Bank of Japan, the Bank of England, the Reserve Bank of Australia, and the Reserve Bank of New Zealand.

The Bank of Canada's Standing Facilities

At the end of each banking day, each LVTS participant must bring its settlement balance with the Bank of Canada close to zero. Of course, it can scarcely be expected that LVTS participants will always be successful in ending up with near-zero settlement balances. The Bank of Canada therefore stands ready (with what we call **standing liquidity facilities**) to lend to or borrow from a participant to bring their settlement balances to zero at the end of the banking day. As already noted, participants also know with certainty the rates applicable to positive and negative settlement balances with the Bank of Canada. (See the FYI box "Monetary Policy Implementation in the LVTS Environment" and Figure 16-1.)

To permit participating financial institutions to adjust positions with each other (i.e., reduce the costs of either positive or negative positions), the LVTS has a pre-settlement trading period of half an hour at the end of the banking day (6:00–6:30 p.m.). The purpose of the pre-settlement trading period is to allow LVTS participants to reduce their LVTS positions (which resulted during the day from their own transactions and those of their clients) at interest rates better than those at the Bank of Canada's standing liquidity facilities—that is, to provide a window for those participants with excess positions to trade with those in deficit, at a better return than can be achieved at the Bank's facilities, to the advantage of both. In fact, the typical bid–ask spread on overnight funds in the interbank market has been less than ⅛%. This is significantly less than the spread of 50 basis points (during normal times) between the rate charged on overdrafts, i_b, and that paid on deposits by the Bank of Canada, $i_b - 0.50$, after settlement of LVTS multilateral positions.

In general, pre-settlement trading among participants will achieve a zero-settlement balance for each participant on wholesale transactions. However, if at the end of the settlement day a participant has a negative balance on the books of the Bank of Canada, the deficit will be financed by a collateralized advance at the bank rate, i_b. Participants with positive settlement balances at the end of the day are paid interest at the bank rate less 50 basis points (i.e., the bottom of the operating band). Hence, as long as the bank rate is set so that the market bid–ask spread is within the operating band, participants will resolve their nonzero settlement balances among themselves rather than through the Bank of Canada's standing facilities. In fact, in a fully competitive market, participants would be expected to trade at the midpoint of the operating band for the overnight interest rate.

Clearly, the LVTS and the Bank of Canada's standing liquidity facilities have been set up in such a way so as to ensure a determinate demand for settlement balances, (in normal times) treating the costs of deficits and surpluses symmetrically. That is, the cost of holding excess settlement balances (an opportunity cost of 25 basis points) equals the cost of holding deficit levels of settlement balances (a premium of 25 basis points for an overdraft loan). These cost incentives are very important as they encourage banks to target zero-settlement balances at the Bank of Canada and in doing so to deal directly with the market rather than to rely on the Bank's automatic standing liquidity facilities.

The Bank of Canada's Implementation of the Operating Band for the Overnight Interest Rate

It is through its lending and taking of deposits from LVTS participants that the Bank of Canada implements its target band for the overnight interest rate. ***If the overnight rate increases toward the upper limit of the operating band, then the Bank will lend at the bank rate to put a ceiling on the overnight rate***. The bank rate is the ceiling on the overnight rate in the money market for LVTS participants, because LVTS participants are unlikely to borrow overnight funds at a higher interest rate, since they can borrow at the bank rate, i_b, from the Bank of Canada.

If the overnight rate declines toward the lower limit of the operating band, then the Bank will accept deposits from LVTS participants at the bank rate less 50 basis points to put a floor on the overnight rate. The bank rate less 50 basis points, $i_b - 0.50$, is the floor on the overnight rate because LVTS participants are unlikely to lend overnight funds at a lower rate, since they can leave funds on deposit at this rate at the Bank of Canada.

16.2 THE MARKET FOR SETTLEMENT BALANCES AND THE CHANNEL/CORRIDOR SYSTEM FOR SETTING THE OVERNIGHT INTEREST RATE

LO 16.2 Illustrate the market for reserves, and demonstrate how changes in monetary policy can affect the equilibrium overnight interest rate.

In Chapter 15, we saw how open market operations (changes in nonborrowed reserves) and Bank of Canada lending (changes in borrowed reserves) affect the balance sheet of the Bank of Canada and the amount of reserves. The market for reserves (settlement balances) is where the overnight interest rate is determined, and this is why we turn to a supply and demand analysis of this market to analyze how the tools of monetary policy affect the overnight interest rate.

Demand and Supply in the Market for Reserves

The analysis of the market for reserves proceeds in a similar fashion to the analysis of the bond market we conducted in Chapter 5. First, we derive demand and supply curves for reserves. Then, the market equilibrium in which the quantity of reserves demanded equals the quantity of reserves supplied determines the level of the overnight interest rate, the interest rate charged on the loans of these reserves.

Demand Curve

To derive the demand curve for reserves, we need to ask what happens to the quantity of reserves demanded by banks, holding everything else constant, as the overnight interest rate, i_{or}, changes. Banks in Canada are no longer required to hold reserves determined through government legislation, as is the case with banks in the Federal Reserve System in the United States and the European Central Bank. Canadian banks, however, hold sufficient cash reserves against their deposit liabilities in order to manage their own short-term liquidity requirements, such as, for example, meeting their depositors' day-to-day demand for cash. We call these reserves *desired reserves*. Desired reserves

equal the desired reserve ratio, r_d, times the amount of deposits on which reserves are desired.

Reserves in excess of the desired amounts are called *excess reserves*. Excess reserves are insurance against deposit outflows, and the cost of holding these excess reserves is their opportunity cost, the interest rate that could have been earned on lending these reserves out, minus the interest rate that is earned on these reserves, i_{er}, which in Canada, during normal times, equals the bank rate less 50 basis points, $i_{er} = i_b - 0.50$.

When the overnight rate is above the rate paid on excess reserves, i_{er}, then as the overnight rate decreases, the opportunity cost of holding excess reserves falls. Holding everything else constant, including the quantity of desired reserves, the quantity of reserves demanded rises. Consequently, the demand curve for reserves, R_d, slopes downward in Figure 16-2 when the overnight rate is above i_{er}. If, however, the overnight rate begins to fall below the interest rate paid on excess reserves i_{er}, banks do not lend in the overnight market at a lower interest rate. Instead, they just keep on adding to their holdings of excess reserves indefinitely. The result is that the demand curve for reserves, R^d, becomes flat (infinitely elastic) at i_{er} in Figure 16-2.

Supply Curve

The supply of reserves, R^s, can be broken up into two components: the amount of reserves that are supplied by the Bank of Canada's open market operations, called *nonborrowed reserves* (*NBR*), and the amount of reserves borrowed from the Bank of Canada, called *borrowed reserves* (*BR*). The primary cost of borrowing from the Bank of Canada is the interest rate charged by the Bank of Canada on these loans—the bank rate, i_b, which is set at a fixed amount above the overnight rate target and thus changes when the target changes.

Mini-lecture

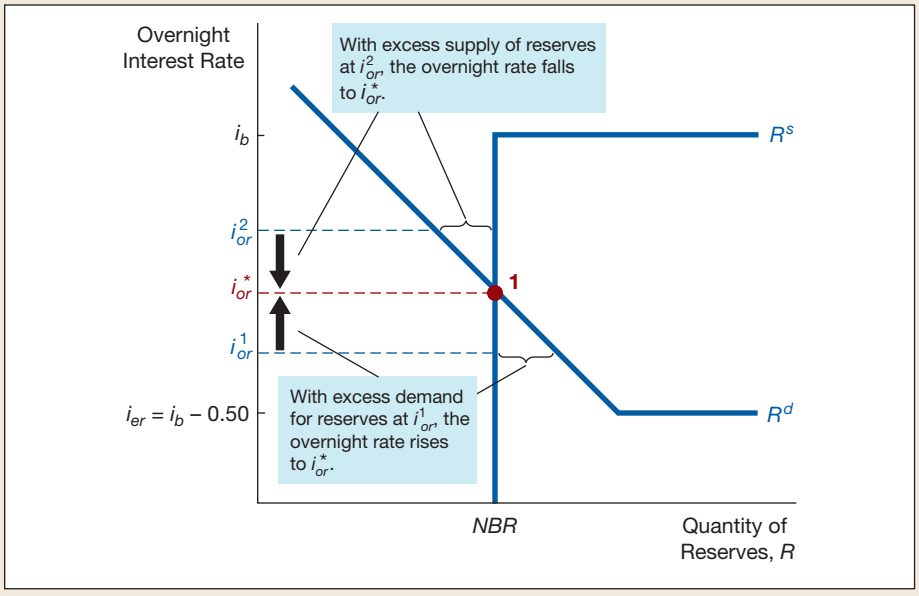


FIGURE 16-2 Equilibrium in the Market for Reserves

Equilibrium occurs at the intersection of the supply curve R^s and the demand curve R^d at point 1 and an interest rate of i^*_{or}.

Because borrowing overnight funds from other banks is a substitute for borrowing (taking out advances) from the Bank of Canada, if the overnight interest rate i_{or} is below the bank rate i_b, then banks will not borrow from the Bank of Canada. Borrowed reserves will be zero because borrowing in the overnight funds market is cheaper. Thus, as long as i_{or} remains below i_b, the supply of reserves will just equal the amount of nonborrowed reserves supplied by the Bank of Canada, *NBR*, and so the supply curve will be vertical, as shown in Figure 16-2. However, as the overnight rate begins to rise above the bank rate, banks will want to keep borrowing more and more at i_b and then lending out the proceeds in the overnight funds market at the higher rate, i_{or}. The result is that the supply curve becomes flat (infinitely elastic) at i_b, as shown in Figure 16-2.

Market Equilibrium

Market equilibrium occurs where the quantity of reserves demanded equals the quantity supplied, $R^d = R^s$. Equilibrium therefore occurs at the intersection of the demand curve R^d and the supply curve R^s at point 1 in Figure 16-2, with an equilibrium overnight interest rate of i^*_{or}. When the overnight interest rate is above the equilibrium rate at i^2_{or}, more reserves are supplied than are demanded (excess supply), and so the overnight interest rate falls to i^*_{or}, as shown by the downward arrow. When the overnight interest rate is below the equilibrium rate at i^1_{or}, more reserves are demanded than are supplied (excess demand), and so the overnight interest rate rises, as shown by the upward arrow.

APPLICATION

How the Bank of Canada's Operating Procedures Limit Fluctuations in the Overnight Interest Rate

An important advantage of the Bank of Canada's current monetary policy procedures is that they limit fluctuations in the overnight interest rate. We can use our supply and demand analysis of the market for reserves to see why.

Suppose that, initially, the equilibrium overnight interest rate is at the overnight rate target of i^*_{or} in Figure 16-3. If the demand for reserves experiences a large unexpected increase, the demand curve would shift to the right to $R^{d'}$, where it now intersects the supply curve for reserves on the flat portion, where the equilibrium overnight rate i^2_{or} equals the bank rate, i_b. No matter how far the demand curve shifts to the right, the equilibrium overnight rate i^2_{or} will just stay at i_b because borrowed reserves will just continue to increase, matching the increase in demand. Similarly, if the demand for reserves experiences a large unexpected decrease, the demand curve shifts to the left to $R^{d'}$, and the supply curve intersects the demand curve on its flat portion, where the equilibrium overnight rate i^1_{or} equals the interest rate paid on excess reserves, $i_{er} = i_b - 0.50$. No matter how far left the demand curve shifts, the equilibrium overnight interest rate i^1_{or} will stay at i_{er} because excess reserves will keep on increasing so that the quantity demanded of reserves equals the quantity of nonborrowed reserves supplied.

Our analysis therefore shows that the ***Bank of Canada's operating procedures limit the fluctuations of the overnight interest rate so that it remains between*** $\boldsymbol{i_{er} = i_b - 0.50}$ ***and*** $\boldsymbol{i_b}$. If the range between i_{er} and i_b is kept narrow enough, then fluctuations around the target rate will be small.

Mini-lecture

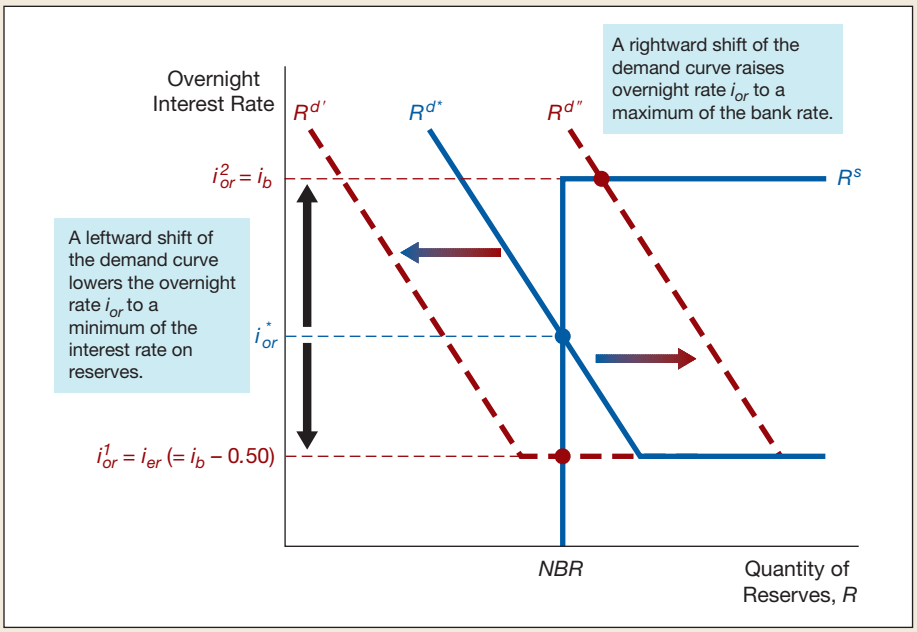

Overnight Interest Rate
$R^{d'}$
R^{d^*}
$R^{d''}$
A rightward shift of the demand curve raises overnight rate i_{or} to a maximum of the bank rate.
$i^2_{or} = i_b$
R^s
A leftward shift of the demand curve lowers the overnight rate i_{or} to a minimum of the interest rate on reserves.
i^*_{or}
$i^1_{or} = i_{er}\ (= i_b - 0.50)$
NBR
Quantity of Reserves, R

FIGURE 16-3 How the Bank of Canada's Operating Procedures Limit Fluctuations in the Overnight Interest Rate

A rightward shift in the demand curve for reserves to $R^{d''}$ will raise the equilibrium overnight rate to a maximum of $i^2_{or} = i_b$, whereas a leftward shift of the demand curve to $R^{d'}$ will lower the overnight rate to a minimum of $i^1_{or} = i_{er}(= i_b - 0.50)$. ◆

16.3 THE BANK OF CANADA'S APPROACH TO MONETARY POLICY

LO 16.3 Discuss the Bank of Canada's approach to monetary policy.

The goal of the Bank of Canada's current monetary policy is to keep the inflation rate within a target range of 1% to 3%, with the midpoint of the inflation target range, 2%, being the most desirable outcome. (See also the FYI box "Price-Level Targeting Versus Inflation-Rate Targeting.") In setting its inflation targets, the Bank of Canada uses the rate of change in the consumer price index (CPI), because it is the most commonly used and understood price measure in Canada. Although the Bank's targets are specified in terms of "headline CPI" (all items), the Bank uses "core CPI," which excludes volatile components such as food, energy, and the effect of indirect taxes, as a measure of the headline rate's trend. Also, defining the inflation targets in terms of ranges provides the Bank of Canada sufficient flexibility to deal with supply shocks beyond those already taken care of by the exclusion of volatile components from core inflation.

Figure 16-4 shows what has happened to the Canadian inflation rate since February 26, 1991, after the Bank's governor and the minister of finance jointly announced

FYI Price-Level Targeting Versus Inflation-Rate Targeting

Currently no central bank is targeting the price level, but some central banks (including the Bank of Canada) are considering whether to replace inflation-rate targeting with price-level targeting.

The advantage of price-level targeting is that it reduces the uncertainty about the long-term purchasing power of money. Consider, for example, the Bank of Canada's objective of keeping the (headline) inflation rate between 1% and 3%, with 2% being the most desirable outcome. This means that over a 30-year period the cumulative inflation rate could be as low as 35% or as high as 143%.

Targeting the price level eliminates that uncertainty. That is, the central bank will make up any lost ground on prices whenever inflation turns out to be less than 2%. Although in theory price-level targeting seems to be superior to inflation-rate targeting, policymakers are reluctant to tinker with the status quo, for now.

a series of declining inflation targets, with a band of plus and minus one percentage point around them. In what follows, we examine the tools used by the Bank of Canada to implement monetary policy, leaving a detailed analysis of the Bank's monetary policy for Chapter 17.

How Monetary Policy Affects the Economy

The Bank of Canada affects interest rates and the level of economic activity by changing the target and operating band for the overnight interest rate. As we saw in Chapter 6, interest rates on different assets tend to move together over time. Hence, changes in the target for the overnight interest rate (and thus the bank rate, i_b) influence other rates, such as the prime rate (the interest rate banks charge to their best customers) and the interest rates on bank deposits and mortgages. These changes in interest rates may also lead to changes in the exchange rate. The level of short-term interest rates and the exchange rate of the Canadian dollar determine the **monetary conditions** in which the economy operates.

The concept of monetary conditions, introduced by the Bank of Canada in its conduct of monetary policy in the early 1990s, focuses on the effect on the economy of both short-term interest rates and the exchange rate. Changes in monetary conditions affect the economy and the Bank's ultimate objective, the inflation rate, only indirectly and are usually felt over a period of several months to several years. This means that the Bank of Canada must always be forward looking in its conduct of monetary policy, anticipating the level of monetary conditions needed today to achieve its ultimate goal of low and stable inflation in the future.

As an example, suppose that the Bank of Canada expects the economy to slow down and wishes to ease monetary conditions. It lowers the target and operating band for the overnight interest rate.[5] As you can see in Figure 16-5, this reduces interest rates and the value of the dollar and leads to an increase in the supply of money, aggregate demand, and the price level, thereby preventing the inflation rate from falling below the lower limit of the 1% to 3% inflation target range.

[5]For example, the Bank might lower the operating band by 25 basis points from 1.5% to 2.0% to 1.25% to 1.75%.

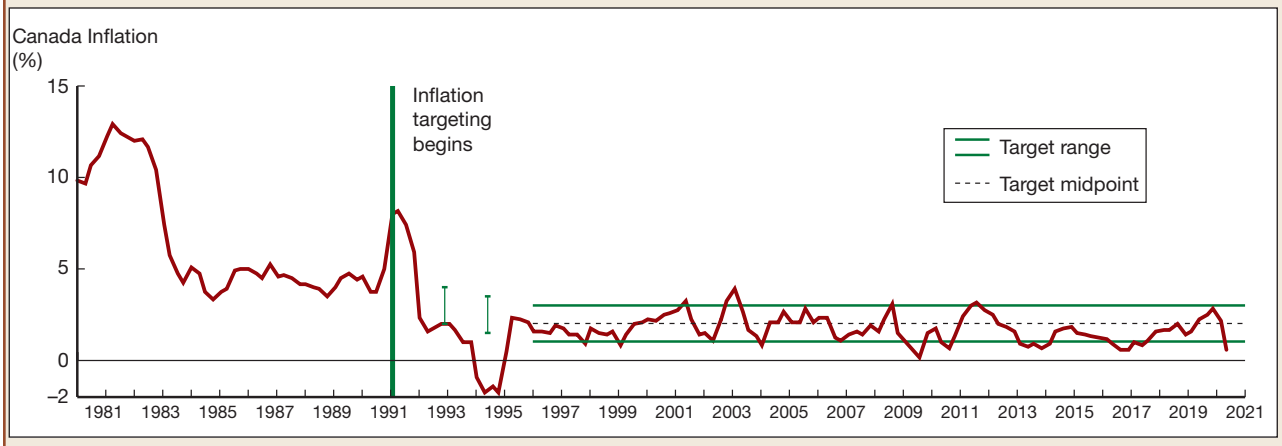


FIGURE 16-4 **Inflation Rates and Inflation Targets for Canada, 1980–2020**

Canada has significantly reduced the rate of inflation and over time has achieved its inflation targets.

Source: Ben S. Bernanke, Thomas Laubach, Frederic S. Mishkin, and Adam S. Poson, *Inflation Targeting: Lessons from the International Experience* (Princeton: Princeton University Press, 1999); and Federal Reserve Bank of St. Louis, FRED database: research.stlouisfed.org/fred2

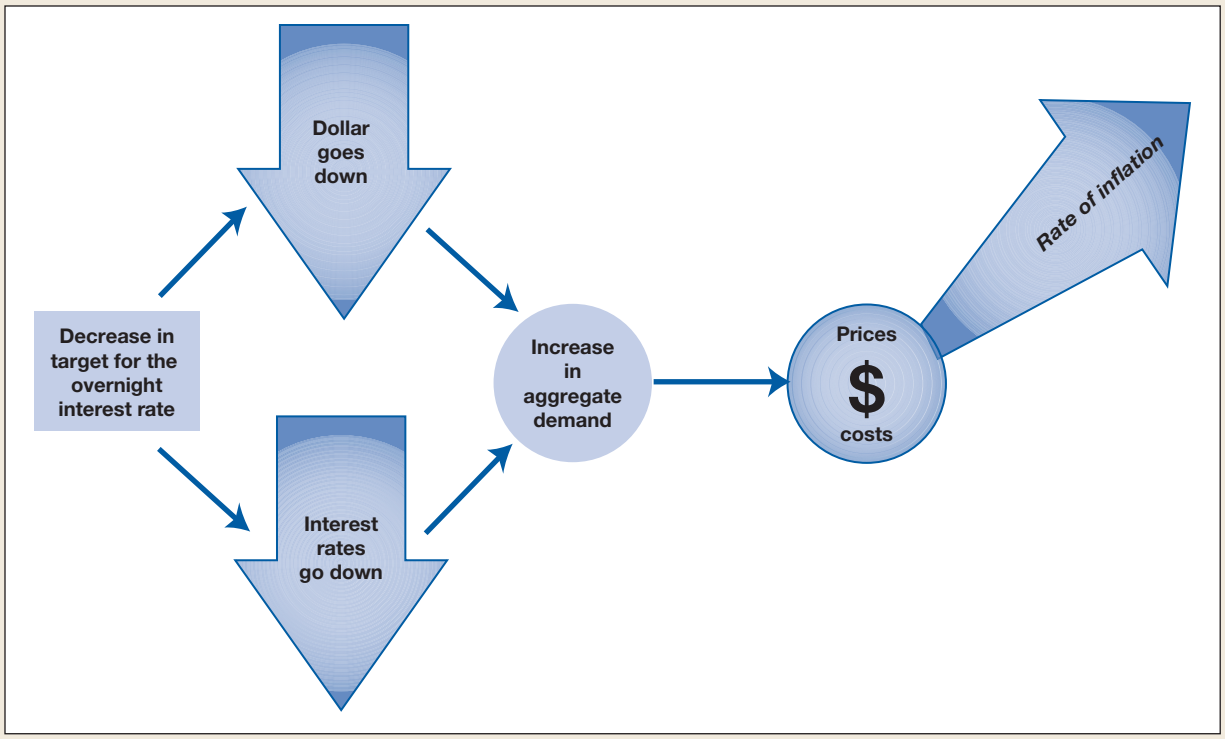


FIGURE 16-5 **How the Bank of Canada Keeps the Rate of Inflation from Falling Below the Lower Limit of the Inflation Target Range**

Source: Bank of Canada website: www.bankofcanada.ca.

In the opposite case, if the Bank expects the economy to be exceeding its capacity at some point in the future, it raises the operating band in order to prevent inflationary pressures from building. The consequent increase in interest rates and the value of the dollar lead to a decline in the supply of money, aggregate demand, and the price level, thereby preventing the inflation rate from moving above the Bank of Canada's upper limit of the 1% to 3% inflation target range—see Figure 16-6.

Hence, by changing the target and operating band for the overnight interest rate, the Bank of Canada sends a signal regarding the direction that it would like interest rates and the money supply to take. A rise in the target and operating band, and thus the bank rate, is a signal that the Bank would like to see higher interest rates and less money in the economy. A fall in the target and operating band is a signal that the Bank would like lower interest rates and more money.

Nominal Interest Rates and Monetary Policy

The Bank of Canada uses the *nominal* overnight interest rate as its operating instrument, but as you will learn in Chapter 24, the effects of monetary policy on economic activity stem from how the *real* interest rate, r, rather than the nominal interest rate, i, affects consumption and investment spending. In fact, the current interest rate targeting approach is based on the belief that it is the real *long-term* interest rate, and not the real short-term interest rate, that affects consumer and business decisions.

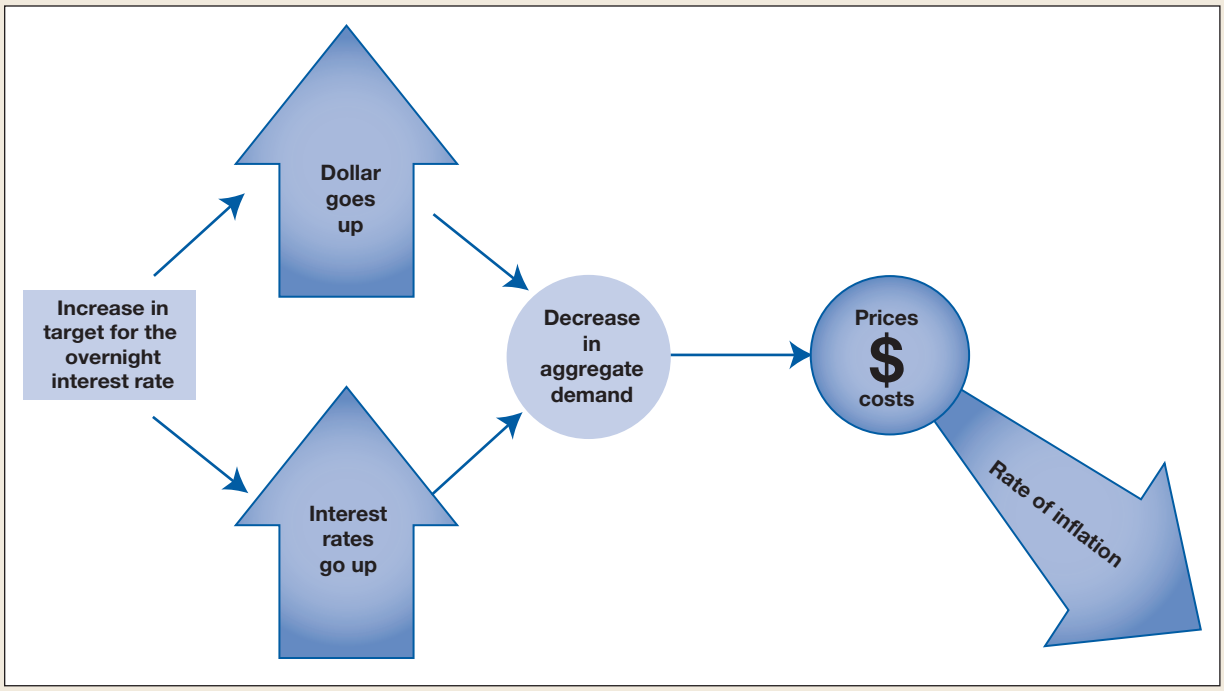


FIGURE 16-6 How the Bank of Canada Keeps the Rate of Inflation from Moving Above the Upper Limit of the Inflation Target Range

Source: Bank of Canada website: www.bankofcanada.ca

How is it that changes in short-term nominal interest rates (such as the Bank of Canada's overnight rate) affect short- and long-term real interest rates? Under the "new Keynesian" assumption of *sticky* prices (that is, under the assumption that prices do not adjust instantaneously to clear the markets), an expansionary monetary policy that lowers the short-term nominal interest rate will also lower the short-term real interest rate, since the aggregate price level adjusts slowly over time and there is not much change in the inflation rate. Moreover, according to the expectations hypothesis of the term structure of interest rates (discussed in detail in Chapter 6), the decline in short-term interest rates will also lead to a decline in long-term interest rates and ultimately to increases in consumer and investment spending.

This transmission mechanism works well even when the short-term nominal interest rate is at or close to zero. With a nominal interest rate of zero, a commitment by the central bank to expansionary monetary policy raises the expected inflation rate π^e, reduces the real interest rate r, and leads to a rise in aggregate output. Thus, expansionary monetary policy could stimulate spending even when the short-term nominal interest rate is at zero. In fact, this mechanism is the key element in monetarist discussions of why an expansionary monetary policy could have prevented the sharp decline in output in the United States during the Great Depression, why it would have helped the Japanese economy when nominal interest rates fell to near zero in the late 1990s, and why it could have helped cope with the global financial crisis and the global economic meltdown.

However, the collapse of stable relationships in financial markets may be causing the term structure of interest rate relationships on which the Keynesian transmission mechanism depends to loosen. For example, the U.S. Federal Reserve raised the target federal funds rate in 17 consecutive meetings from June 2004 to July 2006, from 1% to 5.25%, but long-term interest rates in the United States and around the world declined for most of this period. It has been argued that the reason long-term interest rates did not respond over this period of monetary tightening was the adoption of an inflation-targeting approach to monetary policy by many countries and increased competition from China and India in labour and product markets that contributed to price stability and put downward pressure on long-term interest rates by reducing inflationary expectations.

Similarly, the decline in the federal funds rate from 5.25% in August 2007 to the zero lower bound by late 2008 had not led to the desired decline in long-term interest rates. Why? Because the turmoil in the financial markets had increased risk premiums, thereby keeping long-term nominal interest rates high. Moreover, with collapsing asset prices (stock and real estate prices), there had been an increase in deflationary expectations, which led to a rise in real interest rates. This had the potential to make monetary policy highly contractionary, despite falling short-term nominal interest rates.

16.4 CONVENTIONAL MONETARY POLICY TOOLS

LO 16.4 Summarize how conventional monetary policy tools are implemented and the relative advantages and limitations of each tool.

Now that we understand how the overnight interest rate is determined and the Bank of Canada's approach to monetary policy, we can examine how during normal times the Bank of Canada uses the three tools of monetary policy—open market buyback

operations, settlement balances management, and Bank of Canada lending—to control the money supply and interest rates. These three tools are referred to as **conventional monetary policy**. We will look at each of them in turn to see how the Bank of Canada wields them in practice and how relatively useful each tool is.

Open Market (Buyback) Operations

Open market operations are an important monetary policy tool for many central banks around the world because they are the primary determinants of changes in interest rates and the monetary base, the main source of fluctuations in the money supply. Open market purchases expand bank reserves and the monetary base, thereby lowering short-term interest rates and raising the money supply. Open market sales shrink bank reserves and the monetary base, raising short-term interest rates and lowering the money supply.

The Bank of Canada, however, stopped conducting open market operations in Government of Canada Treasury bills and bonds in 1994, and its most common operations since then have been repurchase transactions with **primary dealers**—the Big Six and the major investment dealers that have signed a Purchase and Resale Agreement with the Bank of Canada. In particular, the Bank uses **special Purchase and Resale Agreements (special PRAs or SPRAs)**, commonly referred to as **repos** or **specials**, as a tool to reduce undesired upward pressure on the overnight interest rate, and **Sale and Repurchase Agreements (SRAs)**, commonly referred to as **reverse repos** or **reverses**, as a tool to reduce undesired downward pressure on the overnight rate.

Let's see how the Bank of Canada uses special PRAs and SRAs in order to reinforce the target for the overnight interest rate during the course of a day.

Special PRAs

Assume that the operating band for the overnight interest rate is 1.5% to 2% and that the Bank of Canada is targeting the overnight rate at the midpoint of the band, at 1.75%. If overnight funds are traded at a rate higher than the target rate of 1.75%, then the Bank of Canada enters into special PRAs, typically at 11:45 a.m., at a price that works out to a 1.75% interest rate, the midpoint of the operating band. That is, the Bank purchases Government of Canada Treasury bills or bonds with an agreement that the seller will repurchase them one business day later. (See Figure 16-7 regarding the mechanics of a special PRA.)

Since the securities are placed with the Canadian Depository for Securities, Canada's central securities depository owned and operated by the financial community, the title of the securities changes hands by electronic instruction. The balance sheets of the Bank of Canada and the primary dealers look like this:

Bank of Canada

Assets		Liabilities	
Government securities	+$100	Settlement balances	+$100

Primary Dealers

Assets		Liabilities	
Settlement balances	+$100	SPRAs	+$100

Hence, ***special Purchase and Resale Agreements (SPRAs) relieve undesired upward pressure on the overnight interest rate***.

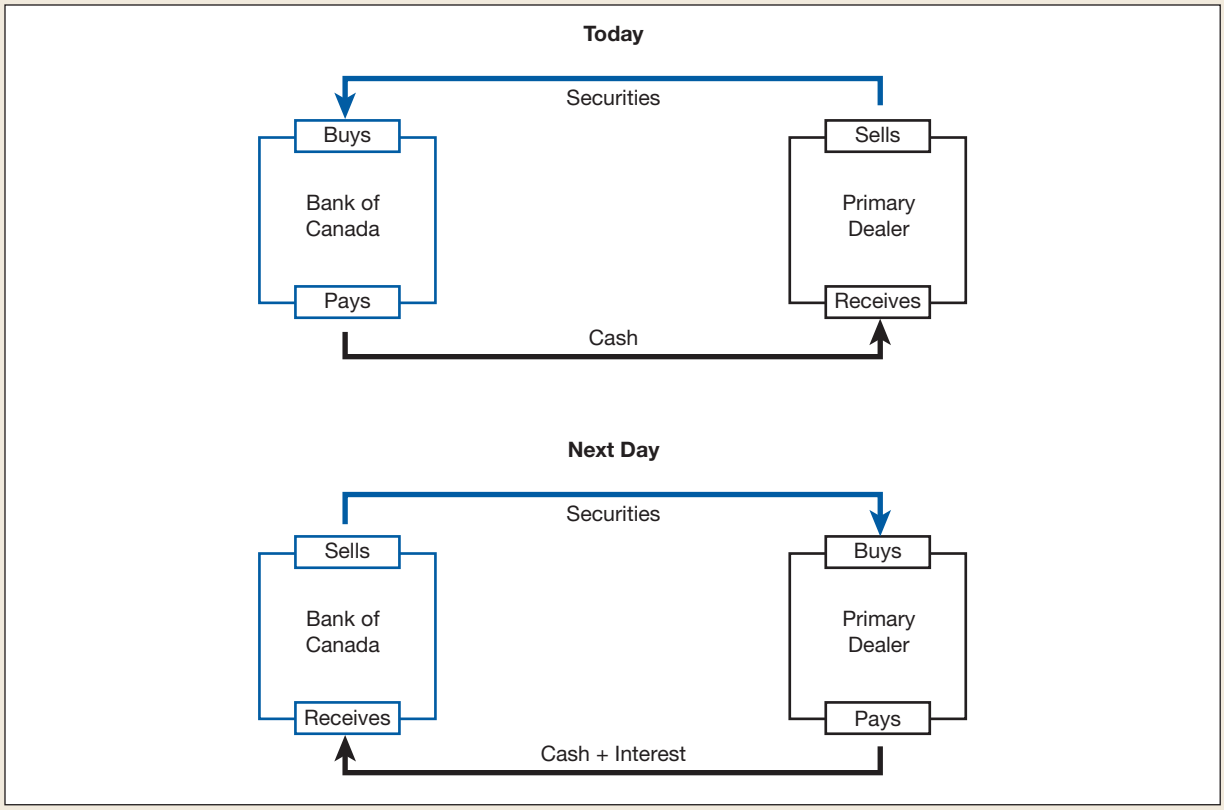


FIGURE 16-7 The Mechanics of a Special PRA Operation

Note: In a special PRA (SPRA), the Bank of Canada buys securities from a primary dealer in exchange for cash, and the primary dealer agrees to buy the securities back (repurchase them) the next day at the original price plus interest, regardless of what happens in the market. In that arrangement, the primary dealer gets an overnight collateralized loan and the Bank of Canada gets some interest.

SRAs If, on the other hand, overnight funds are traded at a rate below the target rate of 1.75%, then the Bank of Canada enters into SRAs, in which the Bank sells government securities and the buyer agrees to sell them back to the Bank one business day later. The balance sheets of the Bank of Canada and the primary dealers now look like this:

Bank of Canada			Primary Dealers		
Assets	**Liabilities**		**Assets**		**Liabilities**
	Settlement balances	−$100	Settlement balances	−$100	
	SRAs	+$100	Government securities	+$100	

Hence, ***Sale and Repurchase Agreements (SRAs) alleviate undesired downward pressure on the overnight financing rate***.

Because the effects on settlement balances of SPRAs and SRAs are reversed on the day the agreement matures, SPRAs and SRAs are actually temporary open market operations. Moreover, as you will see in the next section, because the effects on settlement balances of SPRAs and SRAs are typically neutralized by the end of the day by the Bank of Canada, there is no change at the end of the day in the level of settlement balances in the system.

Advantages of SPRAs and SRAs The Bank of Canada's repurchase transactions, either special PRAs or SRAs, have several advantages over other tools of monetary policy. They occur at the initiative of the Bank of Canada, which has complete control over their volume, and are easily reversed. Moreover, the Bank of Canada can enter into multiple rounds of open market buyback operations and can also conduct these operations outside of the regular timeframe.

Settlement Balances Management

In addition to targeting the overnight interest rate at the midpoint of the operating band, the Bank of Canada also targets the level of settlement balances in the system. Typically, the target level of settlement balances is announced the previous day. Initially, the Bank targeted a daily level of settlement balances of zero, but because of pressures on the overnight interest rate, the Bank made several adjustments to the target level of settlement balances. As of March 19, 2020, the target level of settlement balances is $2 billion.

However, in targeting the overnight interest rate at the midpoint of the operating band, the Bank of Canada usually neutralizes the impact on settlement balances of any open market buyback operations. For example, the Bank neutralizes special PRA operations so as not to leave the system in a surplus position at the end of the day, in which case at least one LVTS participant would end up holding positive settlement balances with the Bank of Canada earning the bank rate less 50 basis points, $i_b - 0.05$. Similarly, the Bank of Canada neutralizes SRA operations so as not to leave the system in a deficit position, in which case at least one LVTS participant would end up holding negative settlement balances with the Bank paying the bank rate, i_b.

Moreover, being the federal government's fiscal agent, the Bank of Canada usually neutralizes the net impact on settlement balances of certain federal government payment flows and Bank of Canada transactions.

The neutralization of the effects of open market buyback operations and of public-sector flows on the level of settlement balances and the adjustment of the level of settlement balances are effected through shifts (transfers) of government deposits (balances) between the government's account at the Bank of Canada and the government's accounts at the LVTS participants. In the pre-LVTS system, government deposit transfers between the government's accounts at banks and its account at the Bank of Canada were effected by the Bank of Canada's daily **drawdowns** (transfers of government deposits from the banks to the Bank of Canada) and **redeposits** (transfers of government deposits from the Bank of Canada to the banks). However, with the instantaneous transfers of funds allowed by the LVTS environment, the drawdown/redeposit mechanism is not a practical tool for government deposit shifting. This is so because the Bank of Canada is not able to debit the government's account with a bank on a day with a net disbursement by the government. In the LVTS environment, the Bank of Canada uses twice-daily

auctions of federal government (Receiver General) balances, the first at 9:15 a.m. (which are collateralized) and the second at 4:15 p.m. (which are uncollateralized), when all the flows affecting the Bank of Canada's balance sheet are known. This neutralizes the effect of market buyback operations and of public-sector flows and also adjusts the level of settlement balances.

Receiver General Auctions To illustrate the process for neutralizing public-sector flows using auctions of Receiver General balances, suppose that there is a net government receipt of $100 (i.e., the government's receipts from the public exceed its payments to the public by $100). Net receipts are drawn on the government's deposits at the LVTS participants, creating claims on the LVTS participants in favour of the Bank of Canada and ultimately reducing settlement balances by an amount equal to the government's net receipts. In the absence of any offsetting transactions by the Bank of Canada, through the workings of supply and demand, a decline in settlement balances will normally cause an increase in the overnight interest rate, as many LVTS participants would have to borrow more to meet their settlement obligations.

However, in targeting the overnight interest rate, the Bank of Canada will neutralize the net government receipt by arranging a net increase of $100 in the government deposit auctions to leave the system unchanged. This procedure results in the following balance sheets for the Bank and the LVTS participants:

Bank of Canada

Assets	Liabilities	
	Government deposits	−100
	Settlement balances	+100

LVTS Participants

Assets		Liabilities	
Settlement balances	+100	Government deposits	+100

Note that government deposits at the Bank of Canada are reduced by $100, and at the same time the settlement balances of the LVTS participants are increased by $100.

If, instead of a net receipt by the government, there were a net disbursement of $100 (i.e., the government's payments to the public exceed its receipts from the public by $100), then the financial system's settlement balances would increase by the same amount. This would prompt a fall in the overnight interest rate, as many LVTS participants would have to borrow less to meet their settlement obligations. The Bank's neutralization process would prevent a decline in the overnight rate by reducing the banking system's settlement balances. This would involve LVTS transfers of $100 (i.e., the difference between the total amount of government balances maturing and the total amount of government balances auctioned) from the government's accounts at the participating institutions to the government's account at the Bank. This procedure results in the following T-accounts:

Bank of Canada

Assets	Liabilities	
	Government deposits	+100
	Settlement balances	−100

LVTS Participants

Assets		Liabilities	
Settlement balances	−100	Government deposits	−100

In this case government deposits at the Bank of Canada are increased by $100, and at the same time the clearing balances of the participants are reduced by $100.

Swaps with the Exchange Fund Account When the Bank of Canada transfers government balances, it usually brings onto its balance sheet Exchange Fund Account assets to back its liabilities. It does so by arranging a swap with the Exchange Fund Account, which holds the country's foreign exchange reserves. This involves a spot purchase and a simultaneous forward sale of foreign exchange. To illustrate the operation, assume that the Bank of Canada temporarily buys $100 of foreign currency assets from the Exchange Fund Account. It credits the government's account on its own books, and the operation results in the following balance sheets for the Bank of Canada and the government:

Bank of Canada			
Assets		**Liabilities**	
Foreign exchange	+100	Government deposits	+100

Government of Canada		
Assets		**Liabilities**
Exchange Account Fund	−100	
Deposits at the Bank of Canada	+100	

We see that government deposit balances at the Bank of Canada increase, and these balances can then be transferred to participants to increase settlement balances, as we saw earlier.

Although the spot transaction adds to the government deposits at the Bank of Canada and enables the Bank to auction government balances, the forward contract between the Bank and the Exchange Fund Account does not affect the settlement balances of participating financial institutions. The Bank simply sells foreign exchange to the Exchange Fund Account in the future at a price agreed upon today. The advantage to the Bank of Canada of using swap transactions with the Exchange Fund Account is that it can bring a temporary change in the level of settlement balances or respond to some event that the Bank thinks will have a significant but not long-lived effect.

As we will learn in Chapter 19, the foreign exchange reserves, held in the Exchange Fund Account, can also be used by the Bank of Canada in international financial transactions to prevent undesirable movements in the exchange rate. For example, a Bank of Canada sale of domestic currency and corresponding purchase of foreign assets in the foreign exchange market leads to a gain in international reserves, an increase in the monetary base and the money supply, and a depreciation of the domestic currency. A Bank purchase of domestic currency and corresponding sale of foreign exchange leads to a loss of international reserves, a decline in the monetary base and the money supply, and an appreciation of the domestic currency. We shall discuss such foreign exchange interventions in detail in Chapter 19.

The Target Level of Settlement Balances In addition to the auctions of Receiver General balances, the Bank of Canada can also adjust the target level of settlement balances if warranted by conditions in the overnight market. That is, the Bank retains the right to adjust the targeted level of settlement balances higher or lower than

the typical setting, depending on pressures on the overnight rate. For example, on January 29, 2016, the Bank of Canada increased the target for the minimum daily level of settlement balances from $150 million to $500 million. Also, on March 19, 2020, the Bank of Canada increased the target level of settlement balances to $2 billion to address the challenges posed by the coronavirus pandemic.

The Bank of Canada, however, is not always able to completely neutralize special PRA and SRA operations and public-sector flows so that the actual level of settlement balances ends up being higher or lower than the target level. However, deviations of the actual level of settlement balances from the target level appear to have insignificant effects on the overnight interest rate the day that they occur, unlike pre-announced changes to the target level of settlement balances, which appear to have a significant effect on the overnight interest rate.

APPLICATION Monetary Control in the Channel/ Corridor System

Suppose that the operating band is 1.5% to 2% and the Bank of Canada expects the economy to exceed its capacity in the near future. To prevent inflationary pressures from building, the Bank wishes to tighten monetary policy by raising the target and operating band by 25 basis points. On one of the eight fixed days for announcing changes to the target and operating band for the overnight rate, the Bank announces at 9:00 a.m. that it is adjusting the target and operating band up from 1.5% to 2% to 1.75% to 2.25%. From this announcement, LVTS participants know that the bank rate shifts from 2% to 2.25%, that the rate on positive settlement balances shifts from 1.5% to 1.75%, and that the Bank of Canada's new target overnight rate, the midpoint of the operating band, shifts from 1.75% to 2%.

Later in the day, at 11:45 a.m., if overnight funds are trading below the target overnight rate (the midpoint of the operating band), the Bank of Canada enters into SRAs to enforce the new target for the overnight rate. That is, it sells government securities to primary dealers, who pay for the securities with settlement balances and agree to sell the securities back to the Bank of Canada on the next business day at a price that works out to an annual interest rate of 2%—the midpoint of the new operating band. Assuming that the Bank of Canada enters into SRAs in the amount of $100, the T-accounts of the Bank of Canada and the primary dealers will be

Bank of Canada

Assets	Liabilities	
	Settlement balances	−100
	SRAs	+100

Primary Dealers

Assets		Liabilities
Settlement balances	−100	
Government securities	+100	

The clearing banks find that their settlement balances have declined by $100, and because they had not been holding any excess reserves, their holdings of settlement

balances are $100 short. Assuming that no other transactions occur during the day, at 4:15 p.m. the Bank of Canada neutralizes the effect on aggregate settlement balances of its issue of SRAs by auctioning off government deposits. The Bank's neutralization process involves auctioning off $100 of government deposits and paying those LVTS participants taking government deposits with settlement balances, bringing the amount of aggregate settlement balances back to zero. The balance sheets of the Bank of Canada and the LVTS participants now look like this:

Bank of Canada

Assets	Liabilities	
	Settlement balances	+100
	Government deposits	−100

LVTS Participants

Assets		Liabilities	
Settlement balances	+100	Government securities	+100

The Bank of Canada's monetary tightening has been effected through the management of settlement balances, during the course of the day, with no change in the aggregate settlement balances of deposit-taking institutions. As already noted, the change in the overnight rate will influence other interest rates and the level of monetary conditions in which the economy operates. ◆

Bank of Canada Lending

The facility at which banks can borrow reserves from the Bank of Canada is called the **standing lending facility**. The easiest way to understand how the Bank of Canada affects the volume of borrowed reserves is by looking at how the standing lending facility operates.

Operation of the Standing Lending Facility

As already noted, the Bank of Canada offers a lending facility, standing ready to lend (upon presentation of suitable collateral) overnight settlement balances to LVTS participants with negative clearing balances at the end of the business day. The lending rate is the bank rate, denoted by i_b in Figure 16-2, and (in normal times) is 25 basis points higher than the target overnight interest rate so as to penalize participants for not using the overnight interbank market. LVTS participants know with certainty the lending rate applicable to their negative settlement balances with the Bank of Canada.

To see how the standing lending facility works, let's see what happens if there is a large increase in the demand for reserves. This situation is analyzed in Figure 16-8. Suppose that initially the demand and supply curves for reserves intersect at point 1 so that the overnight rate is at its target level, i^*_{or}. Now the increase in reserves shifts the demand curve to R^d_2 and the equilibrium moves to point 2. The result is that borrowed reserves increase from zero to BR, and the overnight rate rises to i_b and can rise no further. The standing lending facility has put a ceiling on the overnight interest rate of i_b.

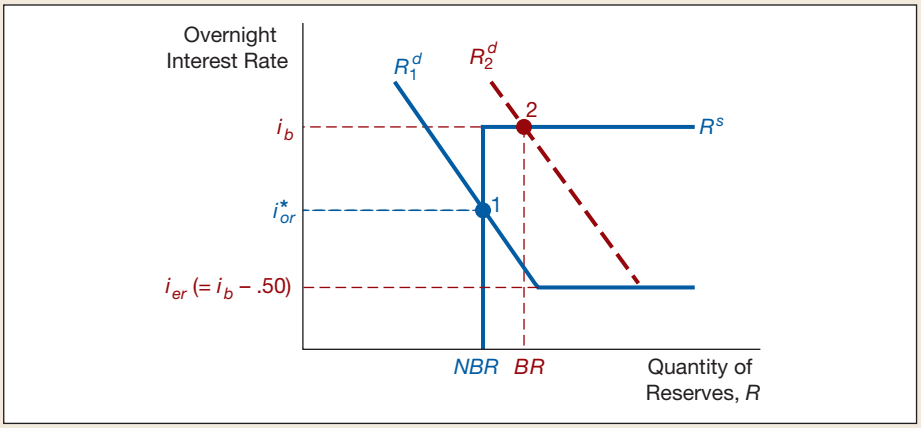


FIGURE 16-8 How the Bank of Canada's Standing Facility Puts a Ceiling on the Overnight Interest Rate

The rightward shift of the demand curve for reserves from R_1^d to R_2^d moves the equilibrium from point 1 to point 2, where $i_{or} = i_b$ and borrowed reserves increase from zero to *BR*.

Lender of Last Resort In addition to its use as a standing liquidity facility to reinforce the operating band for the overnight interest rate, Bank of Canada lending is also important in preventing financial panics. In fact, one of the Bank's most important roles is to be the **lender of last resort** in the Canadian economy.[6] It provides emergency lending assistance (against eligible collateral) for a maximum period of six months (which can be extended for periods up to six months as many times as the Bank judges) to solvent (but illiquid) deposit-taking institutions to prevent bank failures from spinning further out of control, thereby preventing bank and financial panics. The rate that the Bank normally charges on these loans is the bank rate, although the Bank has discretion to charge a higher rate. In providing emergency lending assistance, the Bank always makes a judgment with respect to the trade-off between morally hazardous behaviour and the costs in terms of financial stability. Last-resort lending is a particularly effective way to provide reserves to the banking system during a banking crisis because reserves are immediately channelled to the banks that need them most.

At first glance, it might appear as though the presence of the CDIC, which insures depositors from losses due to a bank's failure up to a limit of $100 000 per account, would make the lender-of-last-resort function of the Bank of Canada superfluous. (The CDIC is described in detail in Chapter 9.) There are two reasons why this is not the case. First, it is important to recognize that the CDIC's insurance fund amounts to a small fraction of the amount of deposits outstanding. If a large number of bank failures occurred, the CDIC would not be able to cover all the depositors' losses. This fact, however, has not weakened the confidence of small depositors in the banking system because the Bank of Canada has been ready to stand behind the banks to provide whatever reserves are needed to prevent bank panics. Second, the large-denomination

[6]Regarding the policy framework that governs the Bank of Canada's lender-of-last-resort activities, see Daniel Fred and Walter Engert, "The Bank of Canada as Lender of Last Resort," *Bank of Canada Review* (Winter 2004–2005).

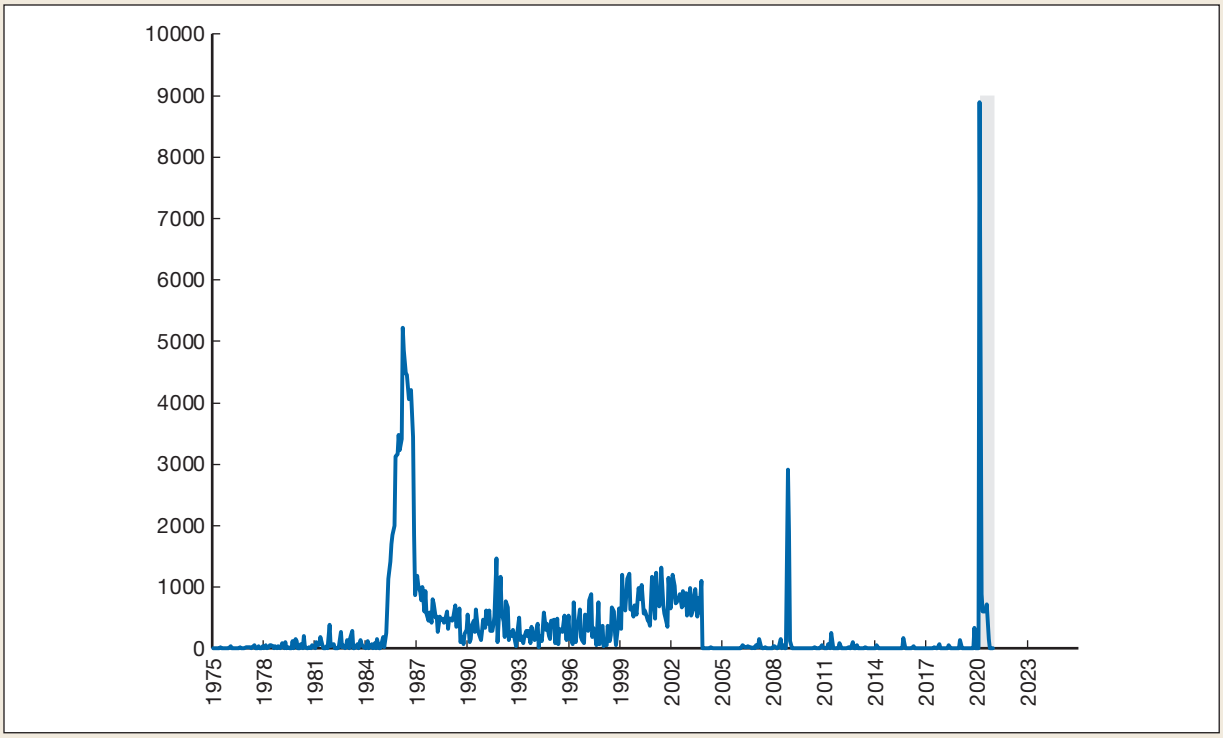


FIGURE 16-9 Bank of Canada Advances to Members of Payments Canada, 1975–2020 (in Millions of Dollars)

Source: Statistics Canada CANSIM series V36663.

deposits in the banking system are not guaranteed by the CDIC because they exceed the $100 000 limit. A loss of confidence in the banking system could still lead to runs on banks from the large-denomination depositors, and bank panics could still occur despite the existence of the CDIC.

The importance of the Bank of Canada's role as lender of last resort is, if anything, more important today because of the 2007–2009 global financial crisis and the 2020 coronavirus crisis. Figure 16-9, which shows Bank of Canada advances to Canadian financial institutions, reveals that the Bank of Canada advanced considerable funds in the recent past to financial institutions facing liquidity crises. Unfortunately, the Bank of Canada's lending policy has not always been successful in preventing financial crises. Two examples of the use of the Bank's lending weapon to avoid bank panics are the provisions of huge loans to the Canadian Commercial Bank and the Northland Bank in 1985. (See the FYI box "Emergency Lending Assistance to Troubled Banks.")

Advantages and Disadvantages of the Bank's Lending Policy The most important advantage of the Bank of Canada's lending policy is that the Bank can use it to perform its role of lender of last resort and stabilize the financial system in times of

stress. Experiences in the 1980s and early 1990s, the Black Monday crash, the terrorist destruction of the World Trade Center, the 2007–2009 global financial crisis, and the Covid-19 crisis indicate that this role has become more important in the past couple of decades.

In the past, Bank of Canada advances were used as a tool of monetary policy, with the bank rate changed to affect interest rates and the money supply. However, because the decisions to take out loans are made by banks and are therefore not completely controlled by the Bank of Canada, while open market buyback operations are completely controlled by the Bank of Canada, the use of lending policy to conduct monetary policy has little to recommend it. This is why the Bank of Canada moved to the current channel/corridor system of monetary control in which the bank rate is not used to set the overnight rate but is only a backup facility to prevent the overnight interest rate from moving outside the operating band.

FYI Emergency Lending Assistance to Troubled Banks

Canadian Commercial and Northland

In 1985, there was public concern over the quality of the assets of two small Alberta-based banks—the Canadian Commercial Bank and the Northland Bank—which had made many bad loans. Larger depositors, whose accounts exceeded the $60 000 limit insured (at that time) by the CDIC, began to withdraw their deposits, and the failure of the banks was imminent. Because the immediate failure of Canadian Commercial and Northland would have had repercussions on other vulnerable banks, the Bank of Canada, under the advice of the Inspector General of Banks (the predecessor of the Office of the Superintendent of Financial Institutions Canada) made extraordinary advances. Total emergency lending assistance amounted to $1.8 billion so that depositors, including the largest, would not suffer any losses.

In doing this, the Bank of Canada was following the precedent established in 1984 by the Federal Reserve's rescue of Continental Illinois National Bank. Continental Illinois had made bad loans (primarily to businesses in the energy industry and to foreign countries), and rumours of financial trouble in early May 1984 caused large depositors to withdraw over US$10 billion of deposits from the bank.

The Federal Deposit Insurance Corporation (FDIC) arranged a rescue effort in July 1984 that culminated in a US$4.5 billion commitment of funds to save the bank. Still, the Fed had to lend Continental Illinois over US$5 billion—making the Bank of Canada's $1.8 billion advances to Canadian Commercial and Northland look like small potatoes! Although Continental Illinois was taken over by the FDIC, the Fed's action prevented further bank failures, and a potential bank panic was averted.

The Bank of Canada, however, was not as successful in preventing a bank crisis in Canada. With the failure of Canadian Commercial and Northland, rumours of financial trouble caused many large depositors to withdraw large deposits from the Bank of British Columbia, Mercantile Bank, and Continental Bank. By the time Mercantile was acquired by the National Bank of Canada, the Bank of British Columbia by the Hongkong Bank of Canada (now HSBC), and Continental by Lloyds Bank of Canada, the Bank of Canada had provided emergency lending assistance of more than $5 billion (see Figure 16-9). The loss of public confidence in the Canadian banking system led to the financial reforms of 1987–1992 and the consolidating of financial institution supervision under the Office of the Superintendent of Financial Institutions Canada.

16.5 NONCONVENTIONAL MONETARY POLICY TOOLS AND QUANTITATIVE EASING IN THE WAKE OF THE GLOBAL FINANCIAL CRISIS AND THE CORONAVIRUS PANDEMIC

LO 16.5 Explain the key monetary policy tools that are used when conventional policy is no longer effective.

In normal times, conventional monetary policy tools, which expand the money supply and lower interest rates, are enough to stabilize the economy. However, when the economy experiences a full-scale financial crisis like the one we recently experienced, conventional monetary policy tools cannot do the job, for two reasons. First, the financial system seizes up to such an extent that it becomes unable to allocate capital to productive uses, and so investment spending and the economy collapse, as discussed in Chapter 11. Second, the negative shock to the economy can lead to the effective-lower-bound problem, in which the central bank is unable to lower its policy rate (the overnight interest rate for the Bank of Canada) much below zero. A negative overnight interest rate well below zero would imply that banks are willing to earn a much lower return by lending in the overnight funds market than they could earn by holding cash in their vaults, with its zero rate of return. For both of these reasons, central banks needed non-interest-rate tools, known as **nonconventional monetary policy tools**, to stimulate the economy in the wake of events such as the global financial crisis and the coronavirus pandemic. These nonconventional monetary policy tools take four forms: (1) liquidity provision, (2) asset purchases, (3) forward guidance, and (4) negative interest rates on bank deposits at a central bank. (See also the FYI box "Helicopter Money" for another potential nonconventional monetary policy tool.)

Liquidity Provision

Because conventional monetary policy actions were not sufficient to heal the financial markets during the recent financial crisis, the Bank of Canada implemented unprecedented increases in its lending facilities to provide liquidity to the financial markets.

1. **Standing Lending Facility Expansion:** At the outset of the crisis on March 4, 2020, the Bank of Canada lowered its policy rate by 50 basis points, from 1.75% to 1.25%, by another 50 basis points to 0.75% on March 13, 2020, and on March 27, 2020, by another 50 basis points to 0.25%, which is what the Bank considers to be its effective lower bound. Since then, the target for the overnight interest rate has been at 0.25%, the bank rate at the Bank of Canada's standing lending facility is at 0.5%, and the deposit rate at the Bank's standing deposit facility is at 0.25%. In doing so, the Bank of Canada also lowered the operating band for the overnight interest rate from 50 basis points to 25 basis points (from ¼% to ½%) and switched from a symmetric system of monetary control to a floor system of monetary control. That is, instead of targeting the overnight interest rate at the midpoint of the operating band, as it does during normal times, the Bank now targets the overnight interest rate at the bottom of the operating band (that is, at the deposit rate), as it did for a brief period in the aftermath of the global financial crisis (from April 1, 2009, to June 1, 2010).

FYI Bank of Canada Programs During the Covid-19 Pandemic

During the Covid-19 pandemic, the Bank of Canada became very creative in assembling a host of new programs to help restore liquidity to different parts of the financial system. The new programs, the dates on which they were created, and their functions are listed below.

Program	Date of Creation	Function
Canada Mortgage Bond Purchase Program	March 17, 2020	To provide the means for financial institutions to renew mortgages, the Bank of Canada buys Canada Mortgage Bonds—bonds issued by financial institutions to finance their mortgage lending—in the secondary market.
Bankers' Acceptance Purchase Facility	March 23, 2020	The Bank of Canada buys Bankers' Acceptances (BAs) in the secondary market to support the continuous function of the BAs market, a key source of funding for small and medium-sized corporations.
Provincial Money Market Purchase Program	March 25, 2020	To support the market for short-term provincial borrowing, the Bank of Canada purchases up to 20% (initially up to 40%) of directly issued provincial money market securities through the primary issuance market.
Standing Term Liquidity Facility	March 31, 2020	Under this facility, the Bank of Canada lends to eligible financial institutions against a broad set of collateral that includes mortgages in order to improve their ability to fund new lending.
Government of Canada Bond Purchase Program	April 1, 2020	To support market functioning and provide monetary stimulus, the Bank of Canada purchases Government of Canada securities in the secondary market of at least $5 billion per week.
Commercial Paper Purchase Program	April 2, 2020	On behalf of the Bank of Canada, TD Asset Management purchases Commercial Paper (CP) to ensure the normal functioning of the CP market, a key source of short-term financing for firms and public authorities.
Contingent Term Repo Facility	April 6, 2020	This is a standing repo facility offering one-month funding to eligible counterparties on a standing, bilateral basis against securities guaranteed by the Government of Canada or a provincial government.
Provincial Bond Purchase Program	May 7, 2020	Managed by BMO Global Asset Management, the Provincial Bond Purchase Program supplements the Provincial Money Market Purchase Program and aims to maintain a well-functioning provincial funding market.
Corporate Bond Purchase Program	May 26, 2020	Managed by TD Asset Management on behalf of the Bank of Canada, this program aims to support the proper functioning of the Canadian corporate debt market.

2. **Standing Term Liquidity Facility:** The Bank of Canada also enhanced its liquidity tools, notably its standing liquidity facilities and term repo operations, to provide ready access to funding to financial institutions. In particular, the Bank expanded the list of eligible financial institutions that can access lending, widened the collateral it accepts to provide lending, and lengthened the term over which it lends to financial institutions. To encourage additional borrowing, on March 31, 2020, the Bank of Canada also established a new *Standing Term Liquidity Facility* to help financial institutions better manage their liquidity risks and continue to provide loans to households and businesses.
3. **New Lending Programs:** The Bank of Canada broadened its provision of liquidity to the financial system well beyond its traditional lending to banking institutions. In particular, the Bank of Canada launched the *Government of Canada Bond Purchase Program* to purchase Government of Canada securities in the secondary market and the *Canada Mortgage Bond Purchase Program* to purchase Canada Mortgage Bonds—bonds issued by financial institutions to finance their mortgage lending—in the secondary market. The Bank also launched the *Bankers' Acceptance Purchase Facility* to support the bankers' acceptances market, the *Commercial Paper Purchase Program* to support the commercial paper market, the *Provincial Money Market Purchase Program* to support the market for short-term provincial borrowing, the *Provincial Bond Purchase Program* to support the provincial bond market, the *Corporate Bond Purchase Program* to support the corporate debt market, and the *Contingent Term Repo Facility* to support the stability of the financial system. These programs are described in more detail in the FYI box "Bank of Canada Programs During the COVID-19 Pandemic."

Large-Scale Asset Purchases

Before the coronavirus pandemic, the Bank of Canada owned about $100 billion worth of Government of Canada bonds, which was about 15% of the total outstanding. In response to the coronavirus recession, on April 1, 2020, the Bank of Canada started buying at least $5 billion per week of Government of Canada bonds and pledged to continue that pace of large-scale asset purchases until the economy recovers. These large-scale asset purchases led to an unprecedented increase in the Bank of Canada's balance sheet (shown in Figure 16-10). This expansion in the Bank of Canada's balance sheet is referred to as **quantitative easing** because, as shown in Chapter 15, it leads to a huge increase in the monetary base. See the FYI box "Quantitative Easing in the Channel/Corridor System: The Floor System of Monetary Control" for a better idea of how quantitative easing works.

Quantitative Easing Versus Credit Easing

As we just learned, the programs introduced by the Bank of Canada in response to the pandemic crisis led to an unprecedented expansion of the Bank of Canada's balance sheet from about $115 billion in March 2020 to over $600 billion by December 2020.

Quantitative easing was introduced by the Bank of Japan in 2001 and has been used extensively by many central banks around the world (including the Federal Reserve, European Central Bank, and the Bank of England) in the aftermath of the global financial crisis of 2007–2009 as well as the recent coronavirus pandemic. For example, in response to the global financial crisis, quantitative easing led to an

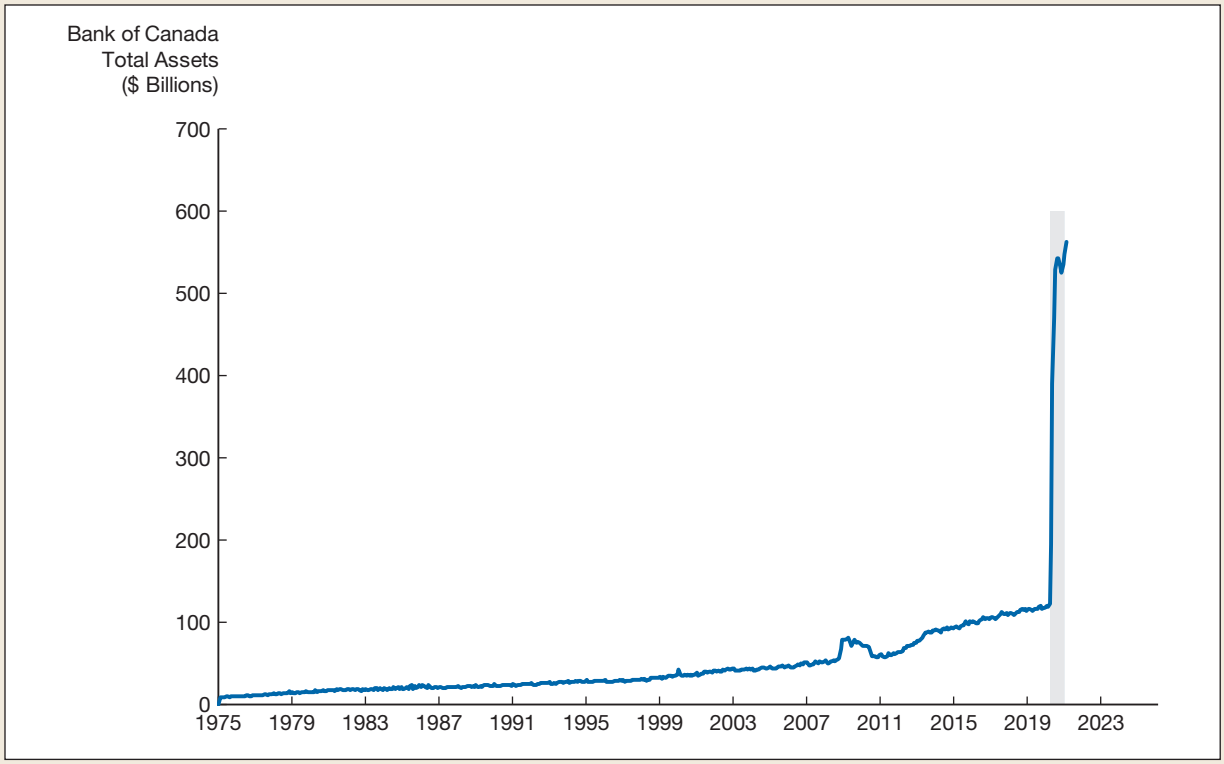


FIGURE 16-10 The Expansion of the Bank of Canada's Balance Sheet, 2001–2021

The size of the Bank of Canada's balance sheet increased significantly during the coronavirus pandemic.

Source: Data from Statistics Canada CANSIM series v36651.

FYI Quantitative Easing in the Channel/Corridor System: The Floor System of Monetary Control

Consider Figure 16-11, which depicts the symmetric corridor system of Figure 16-2, in which the reserve demand is horizontal at $i_b - 0.50$ and the reserve supply is horizontal at i_b, where both rates are tied to the Bank of Canada's target for the overnight interest rate, i^*_{or}. Initially, the supply curve R^s_2 intersects the demand curve R^d in its downward-sloped section at (point 1 and) the target overnight interest rate i^*_{or}.

The central bank can arbitrarily increase the supply of reserves (settlement balances) to undertake expanded asset purchases, shifting its reserve supply schedule from to R^s_1 to R^s_2. But if it goes beyond a certain point (designated 2 in Figure 16-11), then the overnight interest rate will not equal the central bank's interest rate target, i^*_{or}, but rather the interest rate paid on excess reserves, $i_{er} (= i_b - 0.50)$. Hence, the central bank's ability to arbitrarily choose the quantity of reserves comes at the expense of its ability to enforce its target for the overnight interest rate.

Importantly, however, the central bank can independently choose both the interest rate and the quantity of reserves if it targets the overnight interest rate at the excess reserves rate, i_{er}. In this case, the central bank could supply any amount of reserves without any consequence for the policy rate, which would remain unchanged at i_{er}, since it is bounded from below at i_{er}.

This is exactly what the Bank of Canada did from April 1, 2009, to June 1, 2010, when it lowered the band for the overnight interest rate from 50 basis points to 25 basis points (from ½% to ¼%), and instead of targeting the overnight rate at the midpoint of the band (as it does during normal times), it started targeting it at the bottom of the operating band. On June 1, 2010, the Bank re-established the normal operating band of 50 basis points for the overnight interest rate and started again targeting the overnight rate at the midpoint of the band. Then with the coronavirus pandemic, the Bank of Canada reinstituted the floor system of monetary control, with an operating band of 25 basis points (from ¼% to ½%), targeting the overnight interest rate at the bottom of the operating band, at ¼%.

The central bank's ability to arbitrarily choose the quantity of reserves comes at the expense of its ability to enforce its target for the overnight interest rate. If, however, the central bank targets the overnight interest rate at the rate paid on excess reserves, $(i_{er} = i_b - 0.50)$, then it would be able to choose both the overnight interest rate and the quantity of reserves independently.

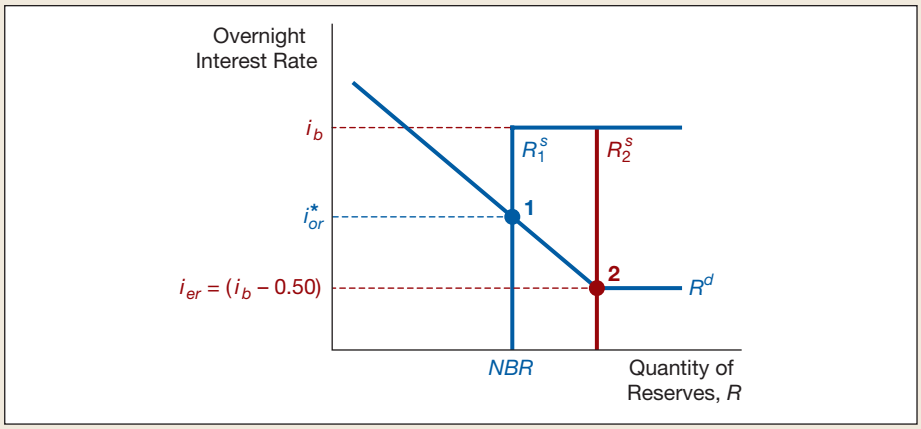


FIGURE 16-11 **Quantitative Easing in the Channel/Corridor System: The Floor System of Monetary Control**

unprecedented expansion of the Federal Reserve's balance sheet from about $900 billion to over $4 trillion by 2014. Then the Fed instituted similar programs in response to the coronavirus pandemic, leading to a further expansion to over $7 trillion in June 2020. Because such an increase in the monetary base would usually result in an expansion of the money supply, it seems as though such an expansion could be a powerful force in stimulating the economy in the near term and possibly producing inflation down the road.

There are reasons to be very skeptical of this hypothesis. First, as we saw in the final application in Chapter 15, the huge expansion in the Fed's balance sheet and the monetary base did not result in a large increase in the money supply, because most of the increase in the monetary base just flowed into holdings of excess reserves. In fact, as shown in Figure 16-12, bank excess reserves in the United States increased to close to $3 trillion by 2014 from around $50 billion before the global financial crisis. Second, because the federal funds rate had already fallen to the effective zero lower bound, the expansion of the balance sheet and the monetary base could not lower short-term interest rates any further and thereby stimulate the economy. Third, an increase in the monetary base does not mean that banks will increase lending, because they can just add to their holdings of excess reserves instead of making loans. Indeed, this appears

FIGURE 16-12 Excess Reserves of Depository Institutions in the United States (in $ Billions)
Source: Federal Reserve Bank of St. Louis FRED database.

to be exactly what happened during the global financial crisis in the United States, when the huge increase in the monetary base led primarily to a massive rise in excess reserves, and bank lending did not increase. A similar phenomenon seems to have occurred when the Bank of Japan engaged in quantitative easing after Japan's stock and real estate market bubbles burst in the 1990s; yet not only did the Japanese economy fail to recover, but inflation even turned negative.

Does skepticism about the merits of quantitative easing mean that the Bank of Canada's nonconventional monetary policy actions during the pandemic were ineffective in stimulating the Canadian economy? The answer is no, because the Bank of Canada's policies are directed not at expanding the Bank of Canada's balance sheet but rather at **credit easing**—that is, altering the composition of the Bank of Canada's balance sheet in order to improve the functioning of particular segments of the credit markets. In this regard, the Federal Reserve chair at the time of the global financial crisis, Ben Bernanke, was adamant that the unprecedented expansion of the Federal Reserve's balance sheet at that time should not be characterized as quantitative easing.

Altering the composition of the central bank's balance sheet can stimulate the economy in several ways. First, when the central bank provides liquidity to a particular segment of the credit markets that has seized up, such liquidity can help unfreeze the market and thereby enable it to allocate capital to productive uses, consequently stimulating the economy. Second, when the central bank purchases particular securities, it increases the demand for those securities, and, as we saw in Chapter 6, such an action can lower the interest rates on those securities relative to rates on other securities. Thus, even if short-term interest rates have hit a floor of zero, asset purchases

can lower interest rates for borrowers in particular credit markets and thereby stimulate spending. For example, purchases of Canada Mortgage Bonds appear to have lowered the interest rates on these securities, which led to a substantial decrease in residential mortgage rates. Purchases of long-term government securities could also lower their interest rates relative to short-term interest rates, and because long-term interest rates are likely to be more relevant to investment decisions, these asset market purchases could boost investment spending. Recent research appears to support this viewpoint.[7]

Forward Guidance

Although short-term interest rates could not be driven below zero in the aftermath of the global financial crisis, central banks had the option of taking a different route in their efforts to achieve lower long-term interest rates, which, as we mentioned above, would stimulate the economy. This route involved a commitment by the central bank to keep the policy rate at zero for a long period of time. To see how this would work, recall our discussion of the expectations theory of the term structure of interest rates in Chapter 6. There we saw that long-term interest rates will equal an average of the short-term interest rates that markets expect to occur over the life of the long-term bond. By committing to the future policy action of keeping the policy rate at zero for an extended period, the central bank could lower the market's expectations of future short-term interest rates, thereby causing the long-term interest rate to fall. Michael Woodford of Columbia University has referred to such a strategy as **management of expectations**, but it is more commonly referred to as **forward guidance**.

The Bank of Canada pursued this forward-guidance strategy when, on April 21, 2009, it lowered the policy rate by 0.25 of a percentage point to 0.25% and committed to holding the policy rate at 0.25% until the second quarter of 2010, conditional on the outlook for inflation. Such a commitment by the Bank of Canada to keep the policy rate at 25 basis points for a long period was expected to shape market expectations of future short-term interest rates and lower long-term interest rates. Also, on July 15, 2020, the Bank of Canada kept the policy rate at 0.25% and announced plans to keep it there until inflation returns to target, which, based on forecasts that the Bank of Canada had separately released, would be after 2022. Although long-term interest rates on Canada securities did subsequently fall, it is not clear how much of this decline was due to the Bank of Canada's forward guidance versus the general weakness of the economy.

There are two types of commitments to future policy actions: conditional and unconditional. The commitment by the Bank of Canada to keep the policy rate at 0.25% until the second quarter of 2010 was *conditional* because it stated that the decision was predicated on a low inflation rate going forward. If economic circumstances changed, the Bank of Canada was indicating that it might abandon the commitment. Alternatively, the Bank of Canada could have made an *unconditional* commitment by just stating that it would keep the policy rate at 0.25% for an extended period without indicating that this decision might change depending on the state of the economy. An unconditional commitment is stronger than a conditional commitment because it does

[7]See, for example, Joseph Gagnon, Mathew Raskin, Julie Remache, and Brian Sack, "Large-Scale Asset Purchases by the Federal Reserve: Did They Work?" *Federal Reserve Bank of New York Economic Policy Review* 17, no. 1 (May 2011): 41–59.

not suggest that the commitment will be abandoned and so is likely to have a larger effect on long-term interest rates. Unfortunately, it has the disadvantage that, even if circumstances change in such a way that it would be better to abandon the commitment, the central bank may feel that it cannot go back on its word and therefore cannot abandon the commitment.

Negative Interest Rates on Banks' Deposits

With inflation very low and their economies weak after the global financial crisis, central banks in Europe and Japan recently began experimenting with a new, nonconventional monetary policy tool, setting interest rates on deposits held by banks at their central banks to be negative. In other words, banks now had to pay their central bank to keep deposits in the central bank. The central bank of Sweden was the first to set negative interest rates on bank deposits in July 2009, followed by the central bank of Denmark in July 2012, the European Central Bank in June 2014, the central bank of Switzerland in December 2014, and the Bank of Japan in January 2016.

Setting negative interest rates on banks' deposits is supposed to work to stimulate the economy by encouraging banks to lend out the deposits they were keeping at the central bank, thereby encouraging households and businesses to spend more. However, there are doubts that negative interest rates on deposits will have the intended expansionary effect.

FYI Helicopter Money

In addition to liquidity provision, asset purchases, forward guidance, and negative interest rate policies, monetary policymakers around the world have also considered (but not yet used) another policy tool in an attempt to increase both output and inflation: helicopter money. The idea comes from Milton Friedman[8] and would consist not of "helicopter drops of money," as Friedman envisioned in a hypothetical scenario, but of the central bank creating money and distributing it directly to the public in the form of fiscal transfer payments such as, for example, an increase in government spending or a tax cut, or more drastically by direct transfers into people's deposit accounts.

In the aftermath of the global financial crisis and the use of quantitative easing, the idea of helicopter money has been taken more seriously by monetary policymakers as a possible tool to stimulate their economies and fight deflationary pressures. Helicopter money is quite similar to quantitative easing, in which central banks create new electronic money to buy assets from financial institutions with the aim of reducing long-term interest rates. The use of helicopter money, however, has the potential to provide stronger stimulus than quantitative easing, because it puts money directly in the hands of consumers, whereas quantitative easing increases bank reserves but not necessarily lending to consumers.

Helicopter money is basically a merging of monetary and fiscal policy, as it represents the creation of new money by the central bank that is distributed in the form of fiscal transfers. The primary argument against the use of helicopter money is that it could reduce the independence of central banks by giving politicians control over the money supply. In this regard, most economists strongly agree that reduced monetary independence would compromise the effectiveness of monetary policy over time and create inflationary pressures.

[8]Milton Friedman, *The Optimum Quantity of Money: And Other Essays* (Chicago: Aldine, 1969).

First, banks might not lend out their deposits at the central bank but instead move them into cash. There would be some cost to doing so because banks would have to build more vaults and hire security guards to protect the cash. Nonetheless, they still might prefer to do this rather than lend them out.

Second, charging banks interest on their deposits might be very costly to banks if they still have to pay positive interest rates to their depositors. In this case, bank profitability would fall. The result might then make banks less likely to lend. So instead of being expansionary, negative interest rates on banks' deposits could cause banks to cut back on lending and therefore be contractionary.

The jury is still out on whether negative interest rates on bank deposits held at central banks will stimulate the economy as intended. Indeed, doubts about the effectiveness of this nonconventional monetary policy tool led Janet Yellen, the chair of the Board of Governors of the Federal Reserve, to announce in 2016 that the Fed was not considering using this tool to stimulate the economy. However, the Federal Reserve's views might change if the U.S. economy becomes much weaker and if the experience with negative interest rates in other countries suggests that this tool is effective in stimulating spending.

16.6 MONETARY POLICY TOOLS OF THE FEDERAL RESERVE AND THE EUROPEAN CENTRAL BANK

LO 16.6 Identify the distinctions and similarities between the monetary policy tools of the Bank of Canada and those of the Federal Reserve and the European Central Bank.

Every central bank practises monetary policy somewhat differently, and in this section we briefly describe the procedures adopted by the Federal Reserve and the European Central Bank. Understanding monetary policy in the United States and the euro area is very important, because it affects interest rates throughout the world.

Policy Tools of the Federal Reserve

In the United States, according to the Depository Institutions Deregulation and Monetary Control Act of 1980, all depository institutions, including commercial banks, savings and loan associations, mutual savings banks, and credit unions, are subject to the same reserve requirements. Up until March 2020, required reserves on all chequable deposits—including non-interest-bearing chequing accounts, NOW accounts, super-NOW accounts, and ATS (automatic transfer savings) accounts—were equal to zero for the first $16.9 million of a bank's chequable deposits, 3% on chequable deposits from $16.9 to $127.5 million, and 10% on chequable deposits over $127.5 million. However, to encourage banks to lend during the coronavirus crisis, on March 15, 2020, the Fed announced a reduction in reserve requirements to zero.

Like the Bank of Canada and most other central banks, the Federal Reserve offers a lending facility, which sets a ceiling on short-term overnight interest rates. The facility at which U.S. banks borrow reserves from the Federal Reserve is called the **discount window**, and the interest rate charged banks for these loans is called the **discount rate**, denoted i_d. However, the primary indicator of the stance of monetary policy in the United States is not the discount rate but the **federal funds rate**, i_{ff}, the interest rate on overnight loans of reserves (known as federal funds) that banks trade among

themselves. In fact, since February 1994, the Fed has announced a federal funds rate target at each Federal Open Market Committee meeting, an announcement that is watched closely by market participants because it affects interest rates throughout the world.

The Fed's discount loans to banks are of three types: primary credit, secondary credit, and seasonal credit. *Primary credit* is the discount lending that plays the most important role in monetary policy. Healthy banks are allowed to borrow all they want at very short maturities (usually overnight) from the primary credit facility, and it is therefore referred to as a standing lending facility. The interest rate on these loans is the discount rate, and it is set higher than the federal funds rate target, usually by 100 basis points (one percentage point) because the Fed prefers that banks borrow from each other in the federal funds market so that they continually monitor each other for credit risk. As a result, in most circumstances the amount of discount lending under the primary credit facility is very small.

Secondary credit is given to banks that are in financial trouble and are experiencing severe liquidity problems. The interest rate on secondary credit is typically set at 50 basis points (0.5 percentage points) above the discount rate. The interest rate on these loans is set at a higher penalty rate to reflect the less-sound condition of these borrowers.

Seasonal credit is given to meet the needs of a limited number of small banks in vacation and agricultural areas that have a seasonal pattern of deposits. The interest rate charged on seasonal credit is tied to the average of the federal funds rate and certificate of deposit rates. The Federal Reserve has questioned the need for the seasonal credit facility because of improvements in credit markets and is thus contemplating eliminating it at some point in the future.

Before the fall of 2008, the Federal Reserve did not pay interest on reserves, but since then it has paid interest on excess reserves at a level that is typically set near the federal funds rate target and therefore changes when the target changes.

The Market for Reserves in the United States The analysis of the market for reserves in the United States proceeds in a similar fashion to the analysis of the market for reserves in Canada depicted in Figure 16-2. The Federal Reserve maintains a two-sided asymmetric corridor system for influencing the federal funds rate as shown in Figure 16-13. The Fed's reserve supply is horizontal at the rate charged by its lending facility, the discount rate i_d, and the bank's reserve demand is horizontal at the interest rate paid by the Fed's deposit facility, i_{er}. Market equilibrium occurs at the intersection of the supply curve R^s and the demand curve R^d at point 1 and a federal funds rate of i^*_{ff}.

Now that we understand how the federal funds rate is determined, we can examine how changes in the four tools of monetary policy—open market operations, discount lending, reserve requirements, and the interest rate paid on excess reserves—affect the market for reserves and the equilibrium federal funds rate in the United States.

Open Market Operations The effect of an open market operation depends on whether the supply curve initially intersects the demand curve in its downward-sloped section or in its flat section. Panel (a) of Figure 16-14 shows what happens if the intersection initially occurs on the downward-sloped section of the demand curve. We have already seen that an open market purchase leads to a greater quantity of reserves supplied; this is true at any given federal funds rate because of the higher amount of

Mini-lecture

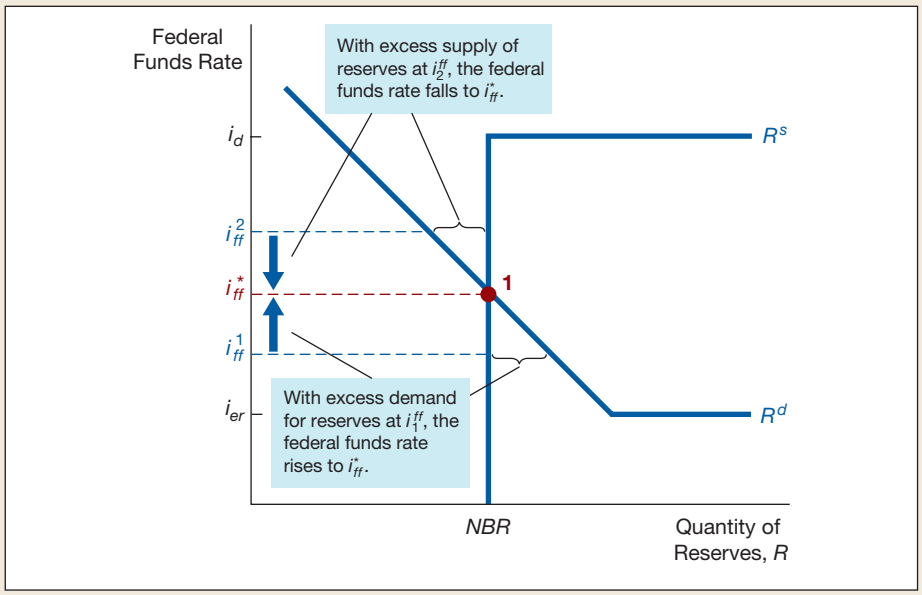


FIGURE 16-13 Equilibrium in the Market for Reserves in the United States
Equilibrium occurs at the intersection of the supply curve R^s and the demand curve R^d at point 1 and an interest rate of i_{ff}^*.

nonborrowed reserves, which rises from NBR_1 to NBR_2. An open market purchase therefore shifts the supply curve to the right from R_1^s to R_2^s and moves the equilibrium from point 1 to point 2, lowering the federal funds rate from i_{ff}^1 to i_{ff}^2.[9] The same reasoning implies that an open market sale decreases the quantity of nonborrowed reserves supplied, shifts the supply curve to the left, and causes the federal funds rate to rise. Because this is the typical situation—since the Fed usually keeps the federal funds rate target above the interest rate paid on reserves—the conclusion is that ***an open market purchase causes the federal funds rate to fall, whereas an open market sale causes the federal funds rate to rise***.

However, if the supply curve initially intersects the demand curve on its flat section, as in panel (b) of Figure 16-14, open market operations have no effect on the federal funds rate. To see this, let's again look at an open market purchase that raises the quantity of reserves supplied, which shifts the supply curve from R_1^s to R_2^s. However, this time we consider the case in which $i_{ff}^1 = i_{er}$ initially. The shift in the supply curve moves the equilibrium from point 1 to point 2, but the federal funds rate remains unchanged at i_{er} because ***the interest rate paid on reserves, i_{er}, sets a floor for the federal funds rate***.

Discount Lending The effect of a discount rate change depends on whether the demand curve intersects the supply curve in its vertical section or its flat section. Panel (a)

[9]We come to the same conclusion if we use the money supply framework outlined in Chapter 15, along with the liquidity preference framework outlined in Chapter 5. An open market purchase raises reserves and the money supply, and then the liquidity preference framework shows that interest rates fall as a result.

Mini-lecture

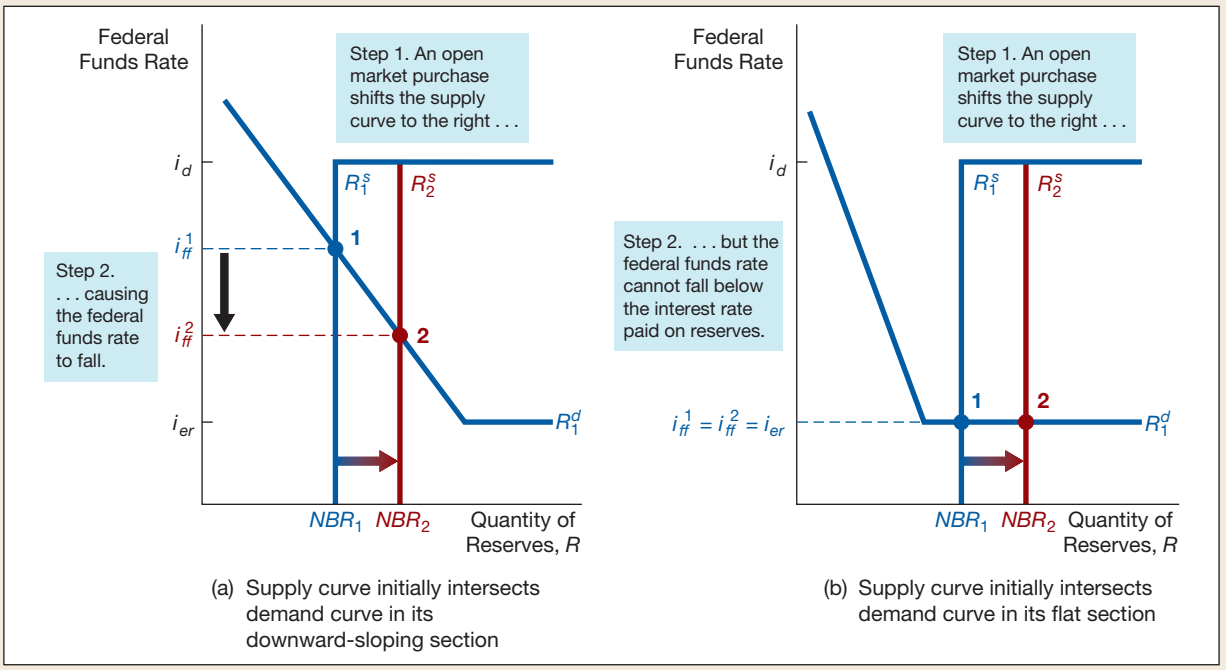

Federal Funds Rate
Step 1. An open market purchase shifts the supply curve to the right . . .
i_d
R_1^s
R_2^s
i_{ff}^1
1
Step 2. . . . causing the federal funds rate to fall.
i_{ff}^2
2
i_{er}
R_1^d
NBR_1
NBR_2
Quantity of Reserves, R
(a) Supply curve initially intersects demand curve in its downward-sloping section
Federal Funds Rate
Step 1. An open market purchase shifts the supply curve to the right . . .
Step 2. . . . but the federal funds rate cannot fall below the interest rate paid on reserves.
$i_{ff}^1 = i_{ff}^2 = i_{er}$
(b) Supply curve initially intersects demand curve in its flat section

FIGURE 16-14 Response to an Open Market Operation

An open market purchase increases nonborrowed reserves, and hence the reserves supplied, and shifts the supply curve from R_1^s to R_2^s. In panel (a), the equilibrium moves from point 1 to point 2, lowering the federal funds rate from to i_{ff}^1 to i_{ff}^2. In panel (b), the equilibrium moves from point 1 to point 2, but the federal funds rate remains unchanged, $i_{ff}^1 = i_{ff}^2 = i_{er}$.

of Figure 16-15 shows what happens if the intersection occurs on the vertical section of the supply curve so that there is no discount lending, and borrowed reserves, BR, are zero. In this case when the discount rate is lowered by the Fed from i_d^1 to i_d^2, the horizontal section of the supply curve falls, as in R_2^s, but the intersection of the supply and demand curves remains at point 1. Thus, in this case, no change occurs in the equilibrium federal funds rate, which remains at i_{ff}^1. Because this is the typical situation—since the Fed now usually keeps the discount rate above its target for the federal funds rate—the conclusion is that ***most changes in the discount rate have no effect on the federal funds rate***.

However, if the demand curve intersects the supply curve on its flat section so that there is some discount lending (i.e., $BR > 0$), as in panel (b) of Figure 16-15, changes in the discount rate do affect the federal funds rate. In this case, initially discount lending is positive and the equilibrium federal funds rate equals the discount rate, $i_{ff}^1 = i_d^1$. When the discount rate is lowered by the Fed from i_d^1 to i_d^2, the horizontal section of the supply curve R_2^s falls, moving the equilibrium from point 1 to point 2, and the equilibrium federal funds rate falls from i_{ff}^1 to $i_{ff}^2 (= i_d^2)$, as shown in panel (b). In this case, BR increases from BR_1 to BR_2.

Reserve Requirements When the required reserve ratio increases, required reserves increase, and hence the quantity of reserves demanded increases for any given

Mini-lecture

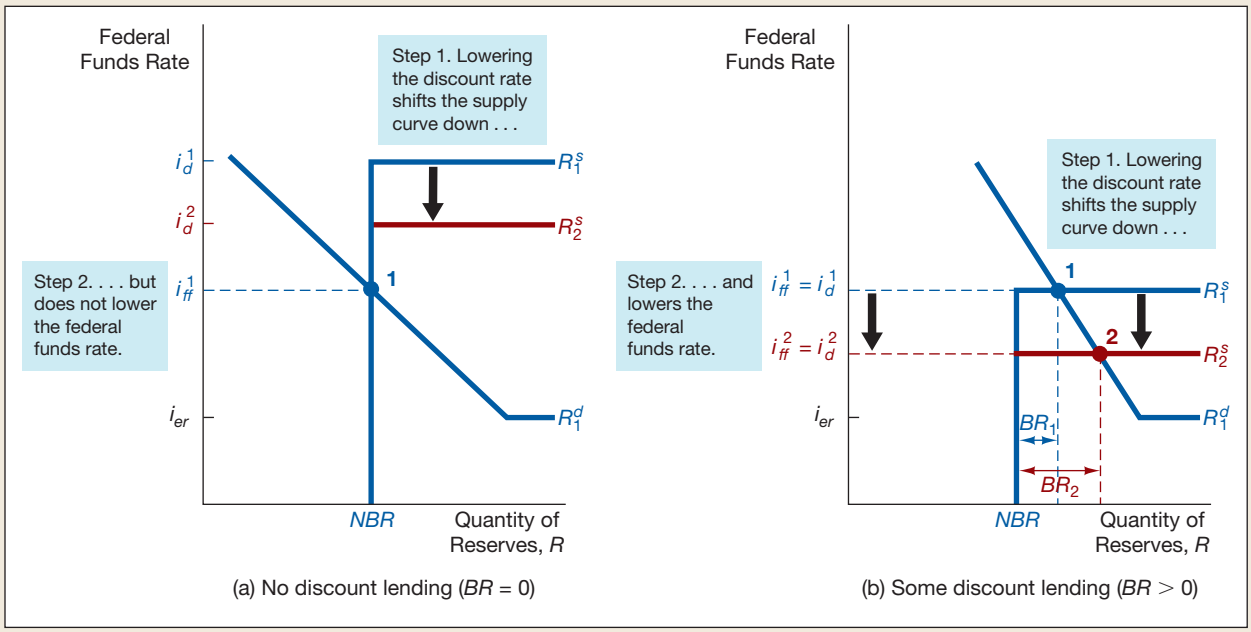


FIGURE 16-15 Response to a Change in the Discount Rate

In panel (a), when the discount rate is lowered by the Fed from i_d^1 to i_d^2, the horizontal section of the supply curve falls, as in R_2^s, and the equilibrium federal funds rate remains unchanged at i_{ff}^1. In panel (b), when the discount rate is lowered by the Fed from i_d^1 to i_d^2, the horizontal section of the supply curve R_2^s falls, and the equilibrium federal funds rate falls from i_{ff}^1 to i_{ff}^2 as borrowed reserves increase.

interest rate. Thus a rise in the required reserve ratio shifts the demand curve to the right from R_1^d to R_2^d in Figure 16-16, moves the equilibrium from point 1 to point 2, and in turn raises the federal funds rate from i_{ff}^1 to i_{ff}^2. The result is that ***when the Fed raises reserve requirements, the federal funds rate rises***.[10]

Similarly, a decline in the required reserve ratio lowers the quantity of reserves demanded, shifts the demand curve to the left, and causes the federal funds rate to fall. ***When the Fed decreases reserve requirements, the federal funds rate falls***.

Interest on Excess Reserves The effect of a change in the interest rate paid by the Fed on excess reserves depends on whether the supply curve intersects the demand curve in its downward-sloping section or its flat section. Panel (a) of Figure 16-17 shows what happens if the intersection occurs on the demand curve's downward-sloping section, where the equilibrium federal funds rate is above the interest rate paid on reserves.

[10]Because an increase in the required reserve ratio means that the same amount of reserves is able to support a smaller amount of deposits, a rise in the required reserve ratio leads to a decline in the money supply. Using the liquidity preference framework from Chapter 5, the fall in the money supply results in a rise in interest rates, yielding the same conclusion given in the text—that raising reserve requirements leads to higher interest rates.

Mini-lecture

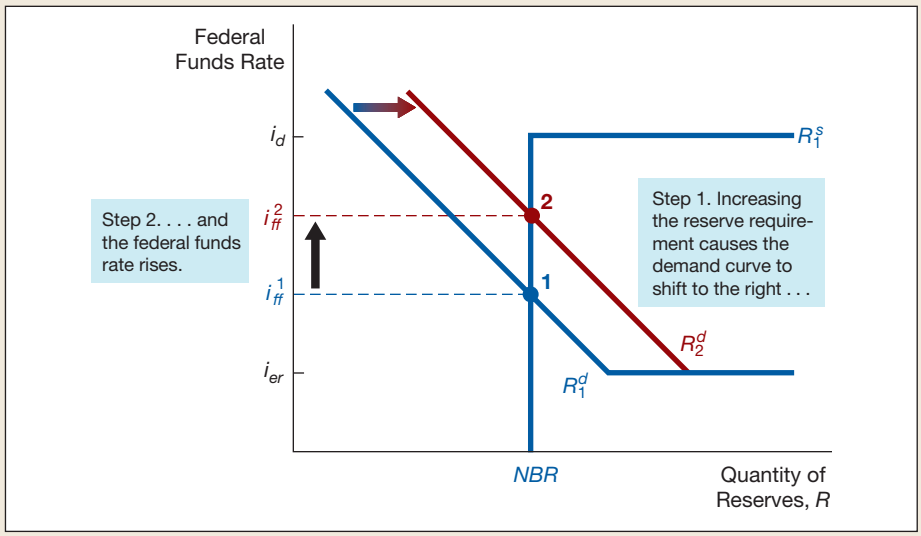


FIGURE 16-16 Response to a Change in Required Reserves
When the Fed raises reserve requirements, required reserves increase, which raises the demand for reserves. The demand curve shifts from R_1^d to R_2^d, the equilibrium moves from point 1 to point 2, and the federal fund rate rises from i_{ff}^1 to i_{ff}^2.

In this case, when the interest rate on reserves is raised from i_{er}^1 to i_{er}^2, the horizontal section of the demand curve rises, as in R_2^d, but the intersection of the supply and demand curves remains at point 1. However, if the supply curve intersects the demand curve on its flat section, where the equilibrium federal funds rate is equal to the interest rate paid on excess reserves, as in panel (b) of Figure 16-17, a rise in the interest rate on excess reserves from i_{er}^1 to i_{er}^2 moves the equilibrium to point 2, where the equilibrium federal funds rate rises from $i_{ff}^1 = i_{er}^1$ to $i_{ff}^2 = i_{er}^2$. ***When the federal funds rate is at the interest rate paid on excess reserves, a rise in the interest rate on excess reserves raises the federal funds rate***.

Monetary Policy Tools of the European Central Bank

Like the Federal Reserve, the European System of Central Banks (usually referred to as the European Central Bank) signals the stance of its monetary policy by setting a **target financing rate**, which in turn sets a target for the **overnight cash rate**. Like the federal funds rate, the overnight cash rate is the interest rate for very short-term interbank loans. The monetary policy tools used by the European Central Bank are similar to those used by the Federal Reserve and involve open market operations, lending to banks, and reserve requirements.

Open Market Operations

Like the Federal Reserve, the European Central Bank uses open market operations as its primary tool for conducting monetary policy and setting the overnight cash rate at the target financing rate. **Main refinancing operations**

Mini-lecture

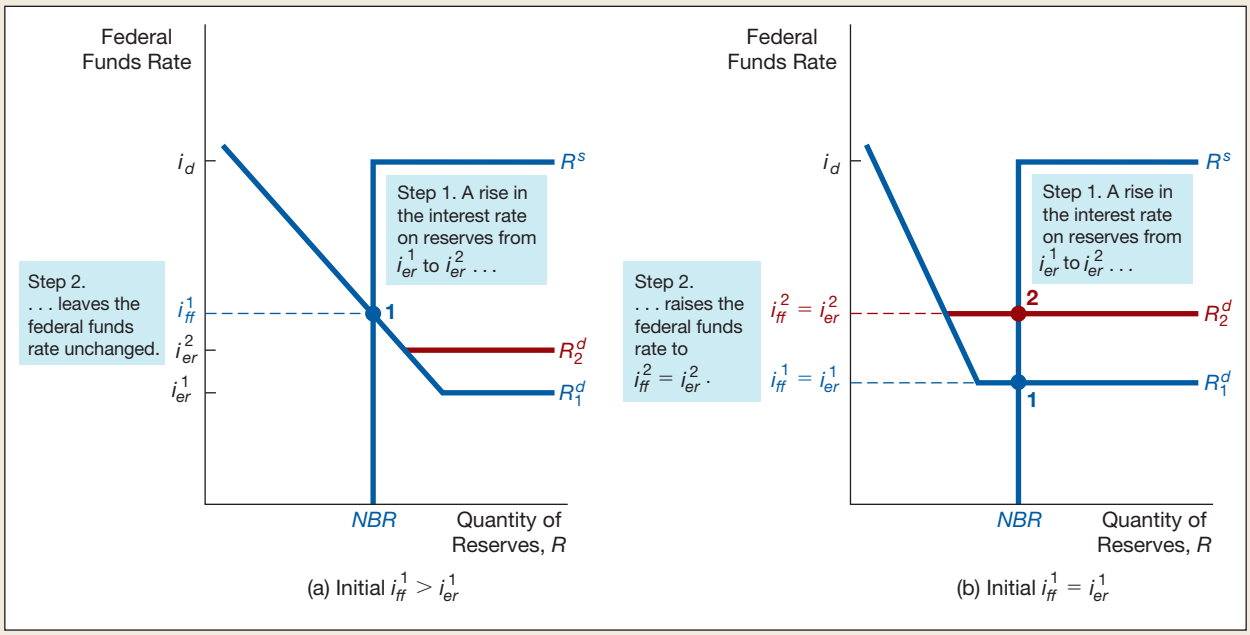


FIGURE 16-17 Response to a Change in the Interest Rate on Reserves

In panel (a), when the equilibrium federal funds rate is above the interest rate paid on reserves, a rise in the interest rate on reserves from i_{er}^1 to i_{er}^2 raises the horizontal section of the demand curve, as in R_1^d, but the equilibrium federal funds rate remains unchanged at i_{ff}^1. In panel (b), when the equilibrium federal funds rate is equal to the interest rate paid on reserves, a rise in the interest rate on reserves from i_{er}^1 to i_{er}^2 raises the equilibrium federal funds rate to $i_{ff}^1 = i_{er}^1$ to $i_{ff}^2 = i_{er}^2$.

are the predominant form of open market operations and are similar to the Fed's repo transactions. They involve weekly **reverse transactions** (purchase or sale of eligible assets under repurchase or credit operations against eligible assets as collateral) that are reversed within two weeks. Credit institutions submit bids, and the European Central Bank decides which bids to accept. Like the Federal Reserve, the European Central Bank accepts the most attractively priced bids and makes purchases or sales up to the point at which the desired amount of reserves is supplied. In contrast to the Federal Reserve, which conducts open market operations in one location at the Federal Reserve Bank of New York, the European Central Bank decentralizes its open market operations by conducting them through the individual national central banks.

A second category of open market operations is the **longer-term refinancing operations**, which are a much smaller source of liquidity for the euro-area banking system and are similar to the Fed's outright purchases or sales of securities. These operations are carried out monthly and typically involve purchases or sales of securities with a maturity of three months. They are not used for signalling the monetary policy stance but instead are aimed at providing euro-area banks access to longer-term funds.

Lending to Banks As is the case for the Bank of Canada and the Fed, the next most important tool of monetary policy for the European Central Bank involves lending to banking institutions, which is carried out by the national central banks, just as discount lending is performed by the individual Federal Reserve Banks in the United States. This lending takes place through a standing lending facility called the **marginal lending facility**. Through these facilities, banks can borrow (against eligible collateral) overnight loans from the national central banks at the **marginal lending rate**, which is set at 100 basis points above the target financing rate. The marginal lending rate provides a ceiling for the overnight market interest rate in the European Monetary Union, just as the discount rate does in the United States.

Interest on Excess Reserves As in the United States, Canada, Australia, and New Zealand, the Eurosystem has another standing facility, the **deposit facility**, in which banks are paid an interest rate on excess reserves that is typically 100 basis points below the target financing rate. The prespecified interest rate on the deposit facility provides a floor for the overnight market interest rate, while the marginal lending rate sets a ceiling. The interest rate on excess reserves set by the European Central Bank is not always positive. As discussed above, the European Central Bank set the interest rate on reserves to negative values starting in June 2014.

Reserve Requirements Like the Federal Reserve, the European Central Bank imposes reserve requirements such that all deposit-taking institutions are required to hold 2% of the total amount of chequing deposits and other short-term deposits in reserve accounts with national central banks. All institutions that are subject to minimum reserve requirements have access to the European Central Bank's standing lending facilities and participate in open market operations.

SUMMARY

1. The Bank of Canada views the overnight interest rate as the centrepiece of its monetary policy implementation. At 9:00 a.m. on the fixed action date, the Bank announces an operating band of 50 basis points for the overnight interest rate. The upper limit of the operating band is the bank rate—the rate the Bank charges LVTS participants that require an overdraft loan to cover negative settlement balances. The lower limit is the rate the Bank pays LVTS participants with positive settlement balances.
2. The Bank of Canada targets the value of the overnight interest rate within its operating band, in normal times at the midpoint of the band. In doing so, the Bank intervenes in the overnight market using open market buyback operations at the target rate. If the overnight rate is trading above the target rate, the Bank uses repos in which it purchases Government of Canada securities from primary dealers with an agreement to resell them on the next business day. If the overnight rate is too low relative to the target rate, the Bank uses reverse repos in which it sells Government of Canada securities to primary dealers with an agreement to buy them back on the next day.
3. The Bank uses government deposit shifting to neutralize public-sector flows that affect LVTS participants' settlement balances—this in effect is a cash setting, a cash setting that is currently $2 billion. Because its holdings of Government of Canada securities are often much smaller than its monetary liabilities, the Bank brings onto its balance sheet Exchange Fund Account assets to back its liabilities. These amounts are adjusted daily, depending on factors such as the level of financial institution borrowings and/or deposits.

4. In neutralizing the effects of open market buyback operations and of public-sector flows on the level of settlement balances, and also in adjusting the level of settlement balances, the Bank of Canada uses transfers of government deposits (balances) between the government's account at the Bank of Canada and the government's accounts at the LVTS participants. These transfers are effected by twice-daily auctions of federal government (Receiver General) balances, the first at 9:15 a.m. (which are collateralized) and the second at 4:15 p.m. (which are uncollateralized).
5. In normal times, the Bank of Canada relies mainly on its conventional monetary policy tools—standing liquidity facilities (lending and deposit facilities), settlement balances management, and lender-of-last-resort arrangements. However, conventional monetary policy tools are no longer effective when the effective-lower-bound problem occurs, in which the central bank is unable to lower short-term interest rates much below zero. In this situation, central banks use nonconventional monetary policy tools, which involve liquidity provision, asset purchases, and forward guidance. Liquidity provision and asset purchases lead to an expansion of the central bank balance sheet, which is referred to as *quantitative easing*. Expansion of the central bank balance sheet by itself is unlikely to have a large impact on the economy, but changing the composition of the balance sheet, which is accomplished by liquidity provisions and asset purchases and is referred to as *credit easing*, can have a large impact by improving the functioning of credit markets.
6. A supply and demand analysis of the market for reserves in the United States yields the following results: When the Fed makes an open market purchase or lowers reserve requirements, the federal funds rate drops. When the Fed makes an open market sale or raises reserve requirements, the federal funds rate rises. Changes in the discount rate and the interest rate paid on excess reserves may also affect the federal funds rate.
7. The monetary policy tools used by the European Central Bank are similar to those used by the Federal Reserve System and involve open market operations, lending to banks, and reserve requirements. Main financing operations—open market operations in repos that are typically reversed within two weeks—are the primary tool to set the overnight cash rate at the target financing rate. The European Central Bank also operates standing lending facilities, which ensure that the overnight cash rate remains within 100 basis points of the target financing rate.

MyLab Economics Make the grade with MyLab Economics: Select end-of-chapter questions can be found on MyLab Economics. You can practise them as often as you want, and most feature step-by-step guided instructions to help you find the right answer.

QUESTIONS

1. If government deposits at the Bank of Canada are predicted to increase, what open market operations could be undertaken to neutralize the effect on settlement balances?
2. During the holiday season, when the public's holdings of currency increases, what open market operations typically occur? Why?
3. If the government has just paid for major computer upgrades, and as a result its deposits with the Bank of Canada fall, what open market operations could be undertaken?
4. "In the LVTS environment, government deposit shifting is affected by auctions of government balances." Discuss.
5. Most open market operations are currently repurchase agreements. What does this tell us about the likely volume of open market operations relative to dynamic open market operations?
6. "The only way that the Bank of Canada can affect the level of advances is by adjusting the bank rate." Is this statement true, false, or uncertain? Explain your answer.
7. If the Bank of Canada did not administer the operating band, what do you predict would happen to the money supply if the bank rate were several percentage points below the overnight rate?
8. Discuss how the operating band affects interest rates and the money supply in the economy.

9. "Last-resort lending is no longer needed because the presence of the CDIC eliminates the possibility of bank panics." Is this statement true, false, or uncertain?
10. The benefits of using last-resort lending to prevent bank panics are straightforward. What are the costs?
11. You often read in the newspaper that the Bank of Canada has just lowered the target overnight rate. Does this signal that the Bank is moving to a more expansionary monetary policy? Why or why not?
12. How can the procyclical movement of interest rates (rising during business cycle expansions and falling during business cycle contractions) lead to a procyclical movement in the money supply as a result of the Bank of Canada's lending policy? Why might this movement of the money supply be undesirable?
13. Explain how SPRAs and SRAs affect the overnight rate.
14. "The channel/corridor system for setting interest rates enables the Bank of Canada to set the overnight rate whatever the demand for reserves, including zero demand." Discuss.
15. Compare the use of open market operations and government deposit shifting to control the money supply, using the following criteria: flexibility, reversibility, effectiveness, and speed of implementation.
16. Following the global financial crisis in 2008, assets on the Federal Reserve's balance sheet increased dramatically, from approximately $800 billion at the end of 2007 to over $4 trillion today. Many of the assets held are longer-term securities acquired through various loan programs instituted as a result of the crisis. In this situation, how could reverse repos (matched sale–purchase transactions) help the Fed reduce its assets held in an orderly fashion while reducing potential inflationary problems in the future?
17. What are the advantages and disadvantages of quantitative easing as an alternative to conventional monetary policy when short-term interest rates are at the zero lower bound?
18. Why is the composition of the central bank's balance sheet a potentially important aspect of monetary policy during an economic crisis?
19. What is the main advantage and the main disadvantage of an unconditional policy commitment?
20. How do the monetary policy tools of the European System of Central Banks compare to the monetary policy tools of the Fed? Does the European Central Bank have a discount lending facility? Does the European Central Bank pay banks an interest rate on their deposits?
21. What is the main rationale behind paying negative interest rates to banks for keeping their deposits at central banks in Sweden, Switzerland, and Japan? What could happen to these economies if banks decide to loan their excess reserves but no good investment opportunities exist?
22. In early 2016, as the Bank of Japan began to push policy interest rates negative, there was a sharp increase in home sales in Japan. Why might this be, and what does it mean for the effectiveness of negative interest rate policy?

APPLIED PROBLEMS

23. If a switch occurs from deposits into currency, what happens to the overnight interest rate? Use the supply and demand analysis of the market for reserves to explain your answer.
24. Using the supply and demand analysis of the market for reserves in the United States, indicate what happens to the federal funds rate, borrowed reserves, and nonborrowed reserves, holding everything else constant, under the following situations:
    a. The economy is surprisingly strong, leading to an increase in the amount of chequable deposits.
    b. Banks expect an unusually large increase in withdrawals from chequing deposit accounts in the future.
    c. The Fed raises the target federal funds rate.
    d. The Fed raises the interest rate on reserves above the current equilibrium federal funds rate.
    e. The Fed reduces reserve requirements.
    f. The Fed reduces reserve requirements and then offsets this action by conducting an open market sale of securities.

DATA ANALYSIS PROBLEMS

1. Go to the CANSIM database and download daily data, from January 1, 1997, to February 6, 2021, for the bank rate (series V39078) and the target overnight rate (series V39079).

a. What is the difference between the two series?

b. Calculate the upper and lower bounds of the operating band. What does the time path of the operating band say about the easiness of monetary policy over this period?

2. **Real-Time Data Analysis** Go to the St. Louis Federal Reserve FRED database and find data on nonborrowed reserves (NONBORRES) and the federal funds rate (FEDFUNDS).

a. Calculate the percent change in nonborrowed reserves and the percentage point change in the federal funds rate for the most recent month of data available and for the same month a year earlier.

b. Is your answer to part (a) consistent with what you expect from the market for reserves in the United States? Why or why not?

17 The Conduct of Monetary Policy: Strategy and Tactics

Learning Objectives

17.1 Define and recognize the importance of a nominal anchor.

17.2 Identify the six potential goals that monetary policymakers may pursue.

17.3 Summarize the distinctions between hierarchical and dual mandates.

17.4 Compare and contrast the advantages and disadvantages of inflation targeting.

17.5 List the four lessons learned from the global financial crisis, and discuss what they mean for inflation targeting.

17.6 Summarize the arguments for and against central bank policy responses to asset-price bubbles.

17.7 Describe and assess the three criteria for choosing a policy instrument.

17.8 Interpret and assess the performance of the Taylor rule as a hypothetical policy instrument for setting the overnight interest rate.

Preview

Getting monetary policy right is crucial to the health of the economy. Overly expansionary monetary policy leads to high inflation, which decreases the efficiency of the economy and hampers economic growth. Monetary policy that is too tight can produce serious recessions in which output falls and unemployment rises. It can also lead to deflation, a fall in the price level such as occurred in the United States during the Great Depression and in Japan more recently. As we saw in Chapter 11, deflation can be especially damaging to an economy because it promotes financial instability and can worsen financial crises.

Now that we understand the tools that central banks such as the Bank of Canada use to conduct monetary policy, we can consider how central banks *should* conduct monetary policy. To explore this subject, we start by looking at the goals of monetary policy and then examine one of the most important strategies for the conduct of monetary policy, inflation targeting. We then discuss tactics—that is, the choice and setting of the monetary policy instrument.

17.1 THE PRICE STABILITY GOAL AND THE NOMINAL ANCHOR

LO 17.1 Define and recognize the importance of a nominal anchor.

Over the past few decades, policymakers throughout the world have become increasingly aware of the social and economic costs of inflation and more concerned with maintaining a stable price level as a goal of economic policy. Indeed, **price stability**, which central bankers define as low and stable inflation, is increasingly viewed as the most important goal of monetary policy. Price stability is desirable because a rising price level (inflation) creates uncertainty in the economy, and that uncertainty might hamper economic growth. For example, when the overall level of prices is changing, the information conveyed by the prices of goods and services is harder to interpret, which complicates decision making for consumers, businesses, and governments, thereby leading to a less efficient financial system.

Not only do public opinion surveys indicate that the public is hostile to inflation, but a growing body of evidence also suggests that inflation leads to lower economic growth. The most extreme example of unstable prices is *hyperinflation*, such as occurred in Argentina, Brazil, Russia, and Zimbabwe in the recent past. Hyperinflation has proved to be very damaging to the workings of an economy.

Inflation also makes it difficult to plan for the future. For example, it is more difficult to decide how much to put aside to provide for a child's college education in an inflationary environment. Furthermore, inflation can strain a country's social fabric: Conflict might result, because each group in the society may compete with other groups to make sure that its income keeps up with the rising level of prices.

The Role of a Nominal Anchor

Because price stability is so crucial to the long-term health of an economy, a central element in successful monetary policy is the use of a **nominal anchor**—a nominal variable, such as the inflation rate or the money supply, that ties down the price level to achieve price stability. Adherence to a nominal anchor that keeps the nominal variable within a narrow range promotes price stability by directly promoting low and stable inflation expectations. A more subtle reason for a nominal anchor's importance is that it can limit the **time-inconsistency problem**, in which monetary policy conducted on a discretionary, day-by-day basis leads to poor long-term outcomes.

The Time-Inconsistency Problem

The time-inconsistency problem is something we deal with continually in everyday life. We often have a plan that we know will produce a good outcome in the long run, but when tomorrow comes, we just can't help ourselves and we renege on our plan because doing so has short-term gains. For example, we make a New Year's resolution to go on a diet, but soon thereafter we can't resist having one more bite of that rocky road ice cream—and then another bite, and then another bite—and the weight begins to pile back on. In other words, we find ourselves unable to *consistently* follow a good plan over *time*; the good plan is said to be *time-inconsistent* and will soon be abandoned.

Monetary policymakers also face the time-inconsistency problem. They are always tempted to pursue a discretionary monetary policy that is more expansionary than firms or people expect because such a policy would boost economic output (or lower unemployment) in the short run. The best policy, however, is *not* to pursue expansionary policy, because decisions about wages and prices reflect workers' and firms' expectations about policy; when they see a central bank pursuing expansionary policy, workers and firms will raise their expectations about inflation, driving wages and prices up. The rise in wages and prices will lead to higher inflation but will not result in higher output on average. (We examine this issue more formally in Chapter 25.)

A central bank will have better inflation performance in the long run if it does not try to surprise people with an unexpectedly expansionary policy but instead keeps inflation under control. However, even if a central bank recognizes that discretionary policy will lead to a poor outcome (high inflation with no gains in output), it still may not be able to pursue the better policy of inflation control because politicians are likely to apply pressure on the central bank to try to boost output with overly expansionary monetary policy.

A clue as to how we should deal with the time-inconsistency problem comes from how-to books on parenting. Parents know that giving in to a child to keep them from acting up will produce a very spoiled child. Nevertheless, when a child throws a tantrum, many parents give them what they want just to shut them up. Because parents don't stick to their "do not give in" plan, the child expects that they will get what they want if they behave badly, so they continue to throw tantrums, over and over again. Parenting books suggest a solution to the time-inconsistency problem (although they don't call it that): Parents should set behaviour rules for their children and stick to them.

A nominal anchor is like a behaviour rule. Just as rules can help prevent the time-inconsistency problem in parenting by helping adults resist pursuing the discretionary policy of giving in, a nominal anchor can help prevent the time-inconsistency problem in monetary policy by providing an expected constraint on discretionary policy.

17.2 OTHER GOALS OF MONETARY POLICY

LO 17.2 Identify the six potential goals that monetary policymakers may pursue.

Although price stability is the primary goal of most central banks, five other goals are continually mentioned by central bank officials when they discuss the objectives of monetary policy: (1) high employment and output stability, (2) economic growth, (3) stability of financial markets, (4) interest-rate stability, and (5) stability in foreign exchange markets.

High Employment and Output Stability

High employment is a worthy goal for two main reasons: (1) the alternative situation—high unemployment—causes much human misery; and (2) when unemployment is high, the economy has both idle workers and idle resources (closed factories and unused equipment), resulting in a loss of output (lower GDP).

Although it is clear that high employment is desirable, how high should it be? At what point can we say that the economy is at full employment? At first, it might seem that full employment is the point at which no worker is out of a job—that is, when unemployment is zero. But this definition ignores the fact that some unemployment, called *frictional unemployment*, which involves searches by workers and firms to find suitable matchups, is beneficial to the economy. For example, a worker who decides to look for a better job might be unemployed for a while during the job search. Workers often decide to leave work temporarily to pursue other activities (raising a family, travel, returning to school), and when they decide to reenter the job market, it may take some time for them to find the right job.

Another reason that unemployment is not zero when the economy is at full employment is *structural unemployment*, a mismatch between job requirements and the skills or availability of local workers. Clearly, this kind of unemployment is undesirable. Nonetheless, it is something that monetary policy can do little about.

This goal of high employment is not an unemployment level of zero but a level above zero that is consistent with full employment, the point at which the demand for labour equals the supply of labour. This level is called the **natural rate of unemployment**.

Although this definition sounds neat and authoritative, it leaves a troublesome question unanswered: What unemployment rate is consistent with full employment? In some cases, it is obvious that the unemployment rate is too high. On the one hand, the unemployment rate in excess of 20% during the Great Depression, for example, was clearly far too high. In the early 1960s, on the other hand, policymakers thought that a reasonable goal for the unemployment rate was 4%, a level that was well below more recent estimates of the natural rate of unemployment of around 6% for that period, with the result that when the unemployment rate fell to 4% in the late 1960s, inflation accelerated. The natural rate of unemployment has fallen since then, and current estimates of the natural rate of unemployment place it around 4%, but even this estimate is subject to much uncertainty and disagreement. It is possible, for example, that appropriate government policy, such as the provision of better

information about job vacancies or job training programs, might decrease the natural rate of unemployment.

The high-employment goal can be thought of in another way. Because the level of unemployment is tied to the level of economic activity in the economy, a particular level of output is produced at the natural rate of unemployment, which naturally enough is referred to as the **natural rate of output** but is more often referred to as **potential output**.

Trying to achieve the goal of high employment thus means that central banks should try to move the level of output toward the natural rate of output. In other words, they should try to stabilize the level of output around its natural rate.

Economic Growth

The goal of steady economic growth is closely related to the high-employment goal because businesses are more likely to invest in capital equipment to increase productivity and economic growth when unemployment is low. Conversely, if unemployment is high and factories are idle, it does not pay for a firm to invest in additional plants and equipment. Although the two goals are closely related, policies can be aimed specifically at promoting economic growth by directly encouraging firms to invest or by encouraging people to save, which provides more funds for firms to invest. In fact, this approach is the stated purpose of *supply-side economics* policies, which are intended to spur economic growth by providing tax incentives for businesses to invest in facilities and equipment and for taxpayers to save more. Active debate continues over the role that monetary policy should play in boosting growth.

Stability of Financial Markets

As our analysis in Chapter 11 showed, financial crises can interfere with the ability of financial markets to channel funds to people with productive investment opportunities and can lead to a sharp contraction in economic activity. The promotion of a more stable financial system, in which financial crises are avoided, is thus an important goal for a central bank. Indeed, as we saw in Chapter 14, the Bank of Canada was created to promote financial stability.

Interest-Rate Stability

Interest-rate stability is desirable because fluctuations in interest rates can create uncertainty in the economy and make it harder to plan for the future. Fluctuations in interest rates that affect consumers' willingness to buy houses, for example, make it more difficult for consumers to decide when to purchase a house and for construction firms to plan how many houses to build. A central bank may also want to reduce upward movements in interest rates for the reasons we discussed in Chapter 14: Upward movements in interest rates generate hostility toward central banks and lead to demands that their power be curtailed.

The stability of financial markets is also fostered by interest-rate stability, because fluctuations in interest rates create great uncertainty for financial institutions. An increase in interest rates produces large capital losses on long-term bonds and mortgages, losses that can cause the failures of the financial institutions holding them.

Stability in Foreign Exchange Markets

With the increasing importance of international trade to the Canadian economy, the value of the dollar relative to other currencies has become a major consideration for the Bank of Canada. A rise in the value of the dollar makes Canadian industries less competitive with those abroad, and a drop in the value of the dollar stimulates inflation in Canada. In addition, preventing large changes in the value of the dollar makes it easier for firms and individuals purchasing or selling goods abroad to plan ahead. Stabilizing extreme movements in the value of the dollar in foreign exchange markets is thus an important goal of monetary policy. In other countries, which are even more dependent on foreign trade, stability in foreign exchange markets takes on even greater importance.

17.3 SHOULD PRICE STABILITY BE THE PRIMARY GOAL OF MONETARY POLICY?

LO 17.3 Summarize the distinctions between hierarchical and dual mandates.

In the long run, no inconsistency exists between the price stability goal and the other goals mentioned earlier. The natural rate of unemployment is not lowered by high inflation, so higher inflation cannot produce lower unemployment or more employment in the long run. In other words, there is no long-run trade-off between inflation and employment. In the long run, price stability promotes economic growth as well as financial and interest-rate stability. Although price stability is consistent with the other goals in the long run, in the short run price stability often conflicts with the goals of output stability and interest-rate stability. For example, when the economy is expanding and unemployment is falling, the economy may become overheated, leading to a rise in inflation. To pursue the price stability goal, a central bank would prevent this overheating by raising interest rates, an action that would initially cause output to fall and increase interest-rate instability. How should a central bank resolve this conflict among goals?

Hierarchical Versus Dual Mandates

Because price stability is crucial to the long-run health of the economy, many countries have decided that price stability should be the primary, long-run goal for central banks. For example, the Maastricht Treaty, which created the European Central Bank, states, "The primary objective of the European System of Central Banks [ESCB] shall be to maintain price stability. Without prejudice to the objective of price stability, the ESCB shall support the general economic policies in the Community," which include objectives such as "a high level of employment" and "sustainable and non-inflationary growth." Mandates of this type, which put the goal of price stability first and then state that other goals can be pursued as long as price stability is achieved, are known as **hierarchical mandates**. They are the directives governing the behaviour of such central banks as the Bank of England, the Bank of Canada, and the Reserve Bank of New Zealand, as well as the European Central Bank.

In contrast, the legislation that defines the mission of the Federal Reserve states, "The Board of Governors of the Federal Reserve System and the Federal Open Market Committee shall maintain long-run growth of the monetary and credit aggregates commensurate with the economy's long-run potential to increase production, so as to promote effectively the goals of maximum employment, stable prices, and moderate long-term

interest rates." Because, as we learned in Chapter 5, long-term interest rates will be very high if inflation is high, this statement in practice is a **dual mandate** to achieve two coequal objectives: price stability and maximum employment (output stability).

Is it better for an economy to operate under a hierarchical mandate or a dual mandate?

Price Stability as the Primary, Long-Run Goal of Monetary Policy

Because no inconsistency exists between achieving price stability in the long run and the natural rate of unemployment, these two types of mandates are not very different *if maximum employment is defined as the natural rate of employment*. In practice, however, a substantial difference between these two mandates might exist because the public and politicians may believe that a hierarchical mandate puts too much emphasis on inflation control and not enough on stabilizing output.

Because low and stable inflation rates promote economic growth, central bankers have come to realize that price stability should be the primary, long-run goal of monetary policy. Nevertheless, because output fluctuations should also be a concern of monetary policy, the goal of price stability should be seen as primary only in the long run. Attempts to keep inflation at the same level in the short run, no matter what else is happening in the economy, are likely to lead to excessive output fluctuations.

As long as price stability is a long-run, but not short-run, goal, central banks can focus on reducing output fluctuations by allowing inflation to deviate from the long-run goal for short periods and, therefore, can operate under a dual mandate. However, if a dual mandate leads a central bank to pursue short-run expansionary policies that increase output and employment without worrying about the long-run consequences for inflation, the time-inconsistency problem may recur. Concerns that a dual mandate might lead to overly expansionary policy is a key reason why central bankers often favour hierarchical mandates in which the pursuit of price stability takes precedence. Hierarchical mandates can also be a problem if they lead to a central bank behaving as what the former governor of the Bank of England, Mervyn King, referred to as an "inflation nutter"—that is, a central bank that focuses solely on inflation control, even in the short run, and so undertakes policies that lead to large output fluctuations. Deciding which type of mandate is better for a central bank ultimately depends on the subtleties of how the mandate will work in practice. Either type of mandate is acceptable as long as it operates to make price stability the primary goal in the long run but not in the short run.

17.4 INFLATION TARGETING

LO 17.4 Compare and contrast the advantages and disadvantages of inflation targeting.

The recognition that price stability should be the primary long-run goal of monetary policy and that a nominal anchor is a valuable tool in helping to achieve this goal has led to a monetary policy strategy known as *inflation targeting*. **Inflation targeting** involves several elements: (1) public announcement of medium-term numerical objectives (targets) for inflation; (2) an institutional commitment to price stability as the primary, long-run goal of monetary policy and a commitment to achieving the inflation goal; (3) an information-inclusive approach in which many variables (not just monetary aggregates) are used in making decisions about monetary policy; (4) increased transparency of the monetary policy strategy through communication with the public

and the markets about the plans and objectives of monetary policymakers; and (5) increased accountability of the central bank for attaining its inflation objectives. New Zealand was the first country to formally adopt inflation targeting in 1990, followed by Canada in 1991, the United Kingdom in 1992, Sweden and Finland in 1993, and Australia and Spain in 1994. Israel, Chile, and Brazil, among other countries, have also adopted a form of inflation targeting.[1]

Inflation Targeting in New Zealand, Canada, and the United Kingdom

We begin our look at inflation targeting with New Zealand, because it was the first country to adopt this strategy. We then go on to examine inflation targeting in Canada and the United Kingdom, which were the next countries to adopt it.

New Zealand As part of a general reform of the government's role in the economy, the New Zealand parliament passed a new Reserve Bank of New Zealand Act in 1989, which became effective on February 1, 1990. In addition to increasing the independence of the central bank, which transformed it from one of the least independent central banks among developed countries to one of the most independent, the act committed the Reserve Bank to a sole objective of price stability. The act stipulated that the minister of finance and the governor of the Reserve Bank should negotiate and make public a Policy Targets Agreement, a statement that sets out the targets by which monetary policy performance will be evaluated, specifying numerical target ranges for inflation and the dates by which they are to be reached. An unusual feature of the New Zealand legislation is that the governor of the Reserve Bank is held highly accountable for the success of monetary policy. If the goals set forth in the Policy Targets Agreement are not satisfied, the governor is subject to dismissal.

The first Policy Targets Agreement, signed by the minister of finance and the governor of the Reserve Bank on March 2, 1990, directed the Reserve Bank to achieve an annual inflation rate within a 3%–5% range. Subsequent agreements lowered the range to 0%–2% until the end of 1996, when the range was changed to 0%–3%. In 2002, the range was changed again, to 1%–3%. Tight monetary policy brought the inflation rate down from above 5% to below 2% by the end of 1992 (see Figure 17-1, panel (a)), but at the cost of a deep recession and a sharp rise in unemployment. Since then, inflation has typically remained within the targeted range. Since 1992, New Zealand's growth rate has generally been high, and unemployment has decreased significantly.

Canada On February 26, 1991, a joint announcement by the minister of finance and the governor of the Bank of Canada established formal inflation targets. The target ranges were to be 2%–4% by the end of 1992, 1.5%–3.5% by June 1994, and 1%–3% by December 1996. After the new government took office in late 1993, the target range was set at 1%–3% for December 1995 until December 1998 and has been kept at this level. Canadian inflation has also fallen dramatically since the adoption of inflation targets, from above 5% in 1991, to a 0% rate in 1995, and to around 2% subsequently (see Figure 17-1, panel (b)). As was the case in New Zealand, however, this decline was

[1]A precursor to the inflation-targeting strategy is monetary targeting. This strategy is discussed in an appendix to this chapter, "Monetary Targeting," which is available **online.** A second appendix provides a history of Bank of Canada policymaking.

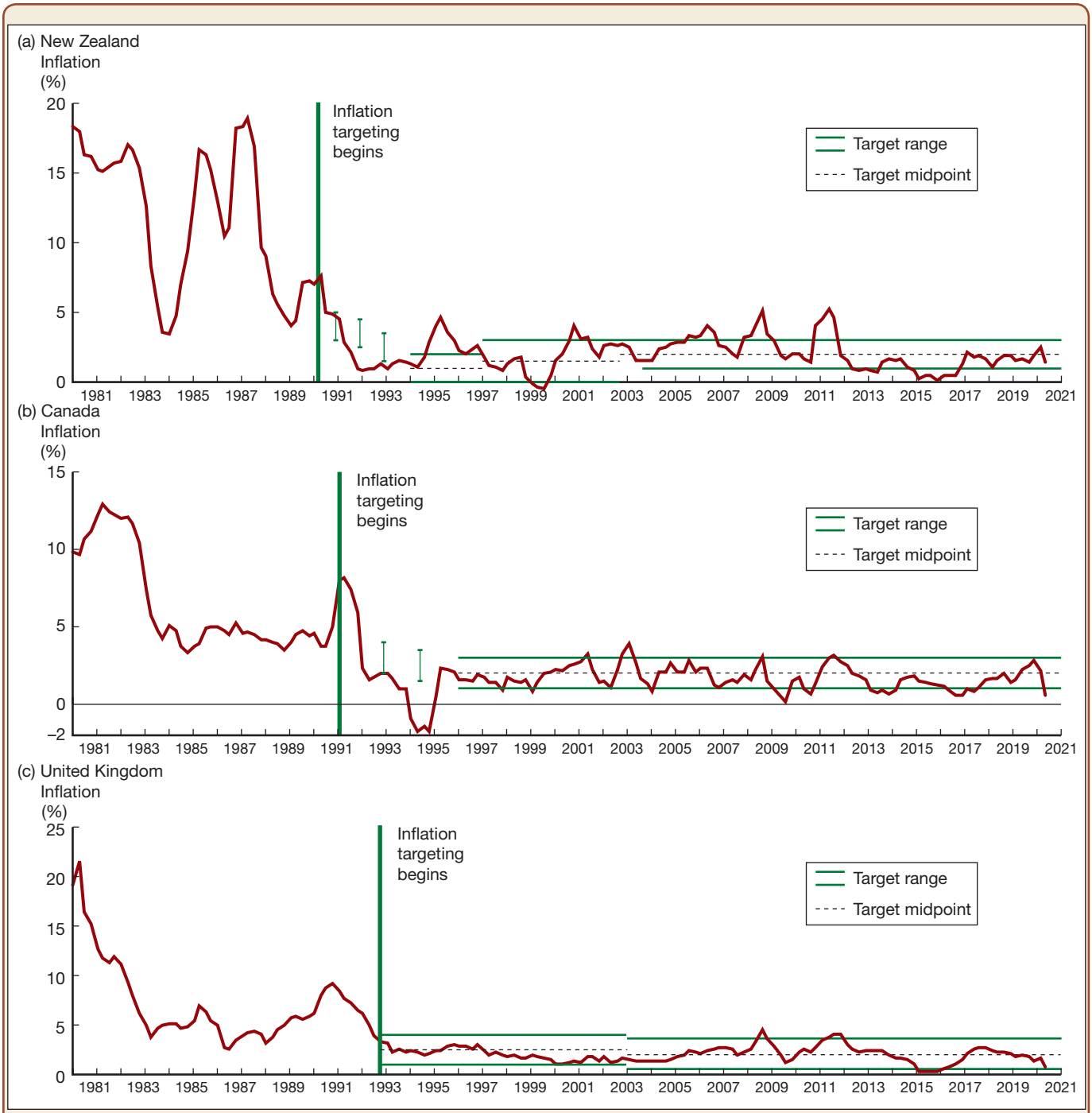


FIGURE 17-1 Inflation Rates and Inflation Targets for New Zealand, Canada, and the United Kingdom, 1980–2020

Inflation-targeting countries have significantly reduced the rate of inflation and over time have achieved their inflation targets.

Sources: Ben S. Bernanke, Thomas Laubach, Frederic S. Mishkin, and Adam S. Poson, *Inflation Targeting: Lessons from the International Experience* (Princeton, NJ: Princeton University Press, 1999); and Federal Reserve Bank of St. Louis FRED database: https://fred.stlouisfed.org/series/NZLCPIALLQINMEI; https://fred.stlouisfed.org/series/CPALCY01CAA661N; https://fred.stlouisfed.org/series/FPCPITOTLZGGBR.

Global The European Central Bank's Monetary Policy Strategy

The European Central Bank (ECB) has been slow to move toward inflation targeting, adopting a hybrid monetary strategy that includes some elements of inflation targeting.* Shortly before the ECB became fully operational in 1999, its governing council defined price stability as an inflation rate below 2%. However, in May 2003, the ECB adopted a goal for inflation over the medium term of "below, but close to, 2%." The ECB's strategy has two key "pillars." First, monetary and credit aggregates are assessed for "their implications for future inflation and economic growth." Second, many other economic variables are used to assess the future economic outlook.

The ECB's strategy is somewhat unclear and has been subject to criticism for this reason. Although the "below, but close to, 2%" goal for inflation sounds like an inflation target, the ECB has repeatedly stated that it does not have an inflation target. This central bank seems to have decided to try to "have its cake and eat it, too" by not committing too strongly to an inflation-targeting strategy. The resulting difficulty in assessing the ECB's strategy has the potential to reduce the accountability of the institution.

*For a description of the ECB's monetary policy strategy, go to the ECB's website at www.ecb.int.

not without cost: Unemployment soared to above 10% from 1991 until 1994 but then declined substantially.

United Kingdom In October 1992, the United Kingdom adopted an inflation target as its nominal anchor, and the Bank of England began to produce an *Inflation Report*, a quarterly report on the progress being made in achieving the target. The inflation target range was initially set at 1%–4% until the next election (spring 1997 at the latest), with the intent that the inflation rate should settle down to the lower half of the range (below 2.5%). In May 1997, the inflation target was set at 2.5% and the Bank of England was given the power to set interest rates henceforth, granting it a more independent role in monetary policy.

Before the adoption of inflation targets, inflation had already been falling in the United Kingdom, with a peak of 9% at the beginning of 1991 and a rate of 4% at the time of adoption (see Figure 17-1, panel (c)). By the third quarter of 1994, it was at 2.2%, within the intended range. Subsequently, inflation rose, climbing slightly above the 2.5% level by the end of 1995, but then fell and has remained close to the target since then. In December 2003, the target was changed to 2.0% for a slightly different measure of inflation. Meanwhile, growth of the U.K. economy has generally been strong, and there has been a substantial decrease in the unemployment rate since the adoption of the inflation target in 1992.[2]

These countries are not the only ones that have moved slowly toward inflation targeting. As the Global box "The European Central Bank's Monetary Policy Strategy" suggests, the European Central Bank also follows a weak form of inflation targeting.

[2]If you are interested in a more detailed discussion of inflation targeting in these and other countries, see Ben S. Bernanke, Thomas Laubach, Frederic S. Mishkin, and Adam S. Posen, *Inflation Targeting: Lessons from the International Experience* (Princeton, NJ: Princeton University Press, 1999).

Advantages of Inflation Targeting

A number of benefits are associated with inflation targeting: reduction of the time-inconsistency problem, increased transparency, increased accountability, consistency with democratic principles, and improved performance.

Reduction of the Time-Inconsistency Problem Because an explicit numerical inflation target increases the accountability of the central bank, inflation targeting can reduce the likelihood that the central bank will fall into the time-inconsistency trap of trying to expand output and employment in the short run by pursuing overly expansionary monetary policy. A key advantage of inflation targeting is that it can help focus the political debate on what a central bank can do in the long run—that is, control inflation—rather than what it cannot do, that is, permanently increase economic growth and the number of jobs through expansionary monetary policy. Thus inflation targeting can reduce political pressures on the central bank to pursue inflationary monetary policy and thereby reduce the likelihood of the time-inconsistency problem.

Increased Transparency Inflation targeting has the advantage that it is readily understood by the public and is thus highly transparent. Indeed, inflation-targeting regimes place great importance on transparency in policymaking and on regular communication with the public. Inflation-targeting central banks have frequent communications with the government, some mandated by law and some in response to informal inquiries, and their officials take every opportunity to make public speeches on their monetary policy strategy. Although these techniques are also commonly used in countries that have not adopted inflation targeting, inflation-targeting central banks have taken public outreach a step further: Not only do they engage in extended public information campaigns, including the distribution of glossy brochures, but they also publish documents like the Bank of England's *Inflation Report*. These documents are particularly noteworthy because, unlike the usual dull-looking, formal reports of central banks, they make use of fancy graphics, boxes, and other eye-catching design elements to engage the public's interest.

The channels of communication just discussed are used by central banks in inflation-targeting countries to explain the following concepts to the general public, financial market participants, and politicians: (1) the goals and limitations of monetary policy, including the rationale for inflation targets; (2) the numerical values of the inflation targets and how they were determined; (3) how the inflation targets are to be achieved, given current economic conditions; and (4) reasons for any deviations from the targets. These communications have improved private-sector planning by reducing uncertainty about monetary policy, interest rates, and inflation; they have promoted public debate of monetary policy, in part by educating the public about what a central bank can and cannot achieve; and they have helped clarify the responsibilities of the central bank and of politicians in the conduct of monetary policy.

Increased Accountability Another key feature of inflation-targeting regimes is the tendency toward increased accountability of the central bank. Indeed, transparency and communication go hand in hand with increased accountability. The strongest case of accountability of a central bank in an inflation-targeting regime is in New Zealand, where the government has the right to dismiss the Reserve Bank's governor if the inflation targets are breached, even for one quarter. In other inflation-targeting countries, the central bank's accountability is less formalized. Nevertheless, the transparency of

policy associated with inflation targeting has tended to make the central bank highly accountable to the public and the government. Sustained success in the conduct of monetary policy as measured against a preannounced and well-defined inflation target can be instrumental in building public support for a central bank's independence and for its policies. This building of public support and accountability occurs even in the absence of a rigidly defined and legalistic standard of performance evaluation and punishment.

Consistency with Democratic Principles Not only is accountability valuable in its own right, but it also makes the institutional framework for the conduct of monetary policy more consistent with democratic principles. The inflation-targeting framework promotes the accountability of the central bank to elected officials, who are given some responsibility for setting the goals of monetary policy and then monitoring the economic outcomes. However, under inflation targeting as it generally has been practised, the central bank has complete control over operational decisions and so can be held accountable for achieving its assigned objectives.

Improved Performance The performance of inflation-targeting regimes has been quite good. Inflation-targeting countries seem to have significantly reduced both the rate of inflation and inflation expectations beyond what likely would have occurred in the absence of inflation targets. Furthermore, once lowered, inflation in these countries has stayed low; following disinflations, the inflation rate in inflation-targeting countries has not bounced back up during subsequent cyclical expansions of the economy.

Disadvantages of Inflation Targeting

Critics of inflation targeting cite four disadvantages of this monetary policy strategy: delayed signalling, too much rigidity, the potential for increased output fluctuations, and low economic growth. We look at each in turn and examine the validity of these criticisms.

Delayed Signalling Inflation is not easily controlled by the monetary authorities. Furthermore, because of the long lags in the effects of monetary policy, inflation outcomes are revealed only after a substantial lag. Thus an inflation target does not send immediate signals to the public and the markets about the stance of monetary policy.

Too Much Rigidity Some economists criticize inflation targeting because they believe it imposes a rigid rule on monetary policymakers and limits their ability to respond to unforeseen circumstances. However, useful policy strategies exist that are "rule-like" in that they involve forward-looking behaviour that limits policymakers from systematically engaging in policies with undesirable long-run consequences. Such policies avoid the time-inconsistency problem and would be best described as "constrained discretion," a term coined by Ben Bernanke and an author of this book.

Indeed, inflation targeting can be described exactly in this way. Inflation targeting, as actually practised, is far from rigid and is better described as "flexible inflation targeting." First, inflation targeting does not prescribe simple and mechanical instructions on how the central bank should conduct monetary policy. Rather, it

requires the central bank to use all available information to determine which policy actions are appropriate to achieve the inflation target. Unlike simple policy rules, inflation targeting never requires the central bank to focus solely on one key variable. Second, inflation targeting, as practised, contains a substantial degree of policy discretion. Inflation targets have been modified depending on economic circumstances, as we have seen. Moreover, central banks under inflation-targeting regimes have left themselves considerable scope to respond to output growth and fluctuations through several devices.

Potential for Increased Output Fluctuations An important criticism of inflation targeting is that a sole focus on inflation may lead to monetary policy that is too tight when inflation is above target and thus may result in larger output fluctuations. Inflation targeting does not, however, require a sole focus on inflation—in fact, experience has shown that inflation targeters display substantial concern about output fluctuations. All the inflation targeters have set their inflation targets above zero.[3] For example, New Zealand, Canada, the United Kingdom, and Sweden currently set the midpoints of their inflation targets at 2%, while Australia sets its midpoint at 2.5%.

The decision by inflation targeters to choose inflation targets above zero reflects the concern of monetary policymakers that particularly low inflation can have substantial negative effects on real economic activity. Deflation (negative inflation, in which the price level actually falls) is especially to be feared because of the possibility that it may promote financial instability and precipitate a severe economic contraction (as discussed in Chapter 11). The deflation in Japan in recent years has been an important factor in the weakening of the Japanese financial system and economy. Inflation-targeting rates above zero make periods of deflation less likely. This is one reason why some economists, both within and outside Japan, called on the Bank of Japan to adopt an inflation target of 2%, which the Bank of Japan finally did in 2013.

Inflation targeting also does not ignore traditional stabilization goals. Central bankers in inflation-targeting countries continue to express their concern about fluctuations in output and employment, and the ability to accommodate short-run stabilization goals to some degree is built into all inflation-targeting regimes. All inflation-targeting countries have been willing to minimize output declines by gradually lowering medium-term inflation targets toward the long-run goal.

Low Economic Growth Another common concern about inflation targeting is that it will lead to low growth in output and employment. Although inflation reduction has been associated with below-normal output during disinflationary phases in inflation-targeting regimes, once low inflation levels were achieved, output and employment returned to levels at least as high as they were before. A conservative conclusion is that once low inflation is achieved, inflation targeting is not harmful to the real economy. Given the strong economic growth after disinflation in many countries (such as New Zealand) that have adopted inflation targets, a case can be made that inflation targeting promotes real economic growth in addition to controlling inflation.

[3]Consumer price indexes have been found to have an upward bias in the measurement of true inflation, so it is not surprising that inflation targets are chosen to exceed zero. However, the actual targets have been set to exceed the estimates of this measurement bias, indicating that inflation targeters have decided on targets for inflation that exceed zero even after measurement bias is accounted for.

17.5 LESSONS FOR MONETARY POLICY STRATEGY FROM THE GLOBAL FINANCIAL CRISIS

LO 17.5 List the four lessons learned from the global financial crisis, and discuss what they mean for inflation targeting.

Our discussion in previous chapters of the course of events that characterized the global financial crisis suggests four basic lessons for economists and policymakers on how the economy works:[4]

1. *Developments in the financial sector have a far greater impact on economic activity than was earlier realized.* Although before the crisis economists and policymakers generally recognized that financial frictions could play an important role in business cycle fluctuations, as we saw in Chapter 11, the global financial crisis made it abundantly clear that the adverse effects of financial disruptions on economic activity could be far worse than originally anticipated.
2. *The effective lower bound on interest rates can be a serious problem.* As we saw in Chapter 16, the effective lower bound on interest rates has forced the Bank of Canada to use nonconventional monetary policy tools during the coronavirus pandemic. Although these nonconventional tools can help stimulate the economy, they are more complicated to use than conventional tools and their impact on the economy is more uncertain, so they may be harder to use effectively.
3. *The cost of cleaning up after a financial crisis is very high.* As we saw in Chapter 11, financial crises are followed by deep recessions. In addition, recoveries from financial crises are very slow. Carmen Reinhart and Vincent Reinhart have documented that economic growth is significantly lower during the decade following financial crises, and unemployment rates stay persistently higher for a decade after crisis episodes. In addition, in the aftermath of financial crises, government indebtedness almost always sharply increases and can lead to defaults on government debt, which have become a major concern in Europe in the aftermath of the most recent crisis.[5]
4. *Price and output stability do not ensure financial stability.* Before the recent financial crisis, the common view in both academia and central banks was that achieving price and output stability would promote financial stability. However, the success of central banks in stabilizing inflation and the decreased volatility of business cycle fluctuations before 2007, which became known as the "Great Moderation," did not protect the economy from financial instability. Indeed, it may have promoted it. The low volatility of both inflation and output fluctuations may have lulled market participants into thinking less risk was present in the economic system than was really the case, leading them to take excessive risks, which in turn helped to fuel the global financial crisis.

[4]For a more detailed discussion of the lessons of the global financial crisis for monetary policy strategy, see Frederic S. Mishkin, "Monetary Policy Strategy: Lessons from the Crisis," in Marek Jarocinski, Frank Smets, and Christian Thimann, eds., *Monetary Policy Revisited: Lessons from the Crisis*, Sixth ECB Central Banking Conference (Frankfurt, Germany: European Central Bank, 2011), 67–118.

[5]See Carmen M. Reinhart and Vincent R. Reinhart, "After the Fall," *Macroeconomic Challenges: The Decade Ahead*, Federal Reserve Bank of Kansas City Economic Symposium, 2010, manuscript available at www.kansascityfed.org; and Carmen M. Reinhart and Kenneth S. Rogoff, *This Time Is Different: Eight Centuries of Financial Folly* (Princeton, NJ: Princeton University Press, 2009).

What implications do these lessons have for monetary policy strategy? We first consider how these lessons might affect our thinking about inflation targeting and then look at how central banks should respond to asset-price bubbles.

Implications for Inflation Targeting

Earlier in the chapter, we outlined the arguments for inflation targeting, and none of the lessons listed above contradicts these arguments. Although support for the inflation-targeting strategy is not weakened by the lessons learned from the financial crisis, these lessons do suggest that inflation targeting may need to be more flexible and also may need to be modified on several dimensions. First, we look at what these lessons might mean for the level of the inflation target.

Level of the Inflation Target As our discussion earlier in the chapter indicated, central banks typically have an inflation target around the 2% level. The seriousness of the effective-lower-bound problem raises the question of whether this target level is too low. A controversial paper written by researchers at the IMF, including its chief economist, Olivier Blanchard, suggested that the inflation target might be raised from the 2% level to the 4% level.[6] The paper argued that with expectations of inflation anchored to this 4% target, by lowering the nominal interest rate to the effective lower bound, say, zero, the real interest rate, $i_r = i - \pi^e$, could be decreased to as low as $-4\%(= 0 - 4\%)$ rather than the rate of $-2\%(= 0 - 2\%)$ associated with the 2% inflation target. Conventional monetary policy, which involves manipulating the nominal policy rate, would then be able to become more expansionary when interest rates fell to the floor of the effective lower bound than it could with the lower inflation target. In other words, the effective lower bound on the policy rate would be less binding with a higher inflation target.

Although this argument is theoretically sound, and raising the inflation target does have its benefits, we also have to look at its costs. The benefits of a higher inflation target accrue only when the effective-lower-bound problem occurs. Although this was a major problem during the global financial crisis, such episodes have not been very frequent. If the effective-lower-bound problem is rare, then the benefits of a higher inflation target are not so large because they are available only infrequently. However, the costs of higher inflation in terms of the distortions it produces in the economy, mentioned early in the chapter, are ongoing. Thus, although these costs may not be that large in any given year, they add up over time and may outweigh the intermittent benefits of a higher inflation target when the effective lower bound occurs.

Another problem with a higher inflation target is that the history of inflation suggests that it is more difficult to stabilize the inflation rate at a 4% level than at a 2% level. Once inflation starts to rise above this level, the public is likely to believe that price stability is no longer a credible goal of the central bank. The question then arises, if a 4% level of inflation is okay, then why not 6%, or 8%, and so on? Indeed, this is what seems to have happened in the 1960s, when economists such as Paul Samuelson and Robert Solow of MIT, both eventual recipients of the Nobel Prize, argued that policymakers should be willing to tolerate higher inflation rates in the 4%–5% range. But when inflation rose to that level, the policy authorities could not contain it at that level, and it kept on rising to double-digit levels by the early 1980s. Getting inflation back down again during the Volcker era was very costly. No central banker wants to go through that cycle again, and this is why central bankers have been so hostile to the IMF researchers' suggestion.

[6]Olivier Blanchard, Giovanni Dell'Ariccia, and Paolo Mauro, "Rethinking Monetary Policy," *Journal of Money, Credit and Banking* 42, no. S1 (September 2010): 199–217.

Although central bankers have not supported raising the inflation target above the 2% level, the effective-lower-bound problem suggests that allowing inflation expectations to fall below 2% could be very costly: It would raise the real interest rate when interest rates hit the effective lower bound after there is a negative shock to the economy.

In this regard, even though the Federal Reserve adopted a 2% inflation objective in January 2012, inflation had been running below 2% for a number of years afterward, so the Fed became concerned that the persistent inflation undershoots of the 2% target could lead to inflation expectations declining below 2%. As a result, in August 2020, after the Fed had conducted a study of its monetary policy strategy, Jay Powell announced that the Fed was modifying its inflation target to be a 2% *average* rather than a 2% annual target. How this modification changes the Fed's monetary policy strategy is described in the FYI box "The Fed's New Monetary Policy Strategy: Average Inflation Targeting."

Flexibility of Inflation Targeting We have seen that inflation targeting as actually practised would be better described as "flexible inflation targeting." However, before the global financial crisis, this flexibility involved allowing some short-run deviations of inflation from the inflation target to promote output stability as well as price stability. Two lessons from the crisis—that financial instability can have devastating effects on the economy and that achieving price and output stability does not ensure financial stability—have led to a recognition that central banks need to pay more attention to financial stability, not only in designing inflation-targeting regimes but also in any monetary policy framework. Particularly important in this regard is the issue of how central banks should respond to asset-price bubbles, the topic we discuss next.

FYI The Fed's New Monetary Policy Strategy: Average Inflation Targeting

The Federal Reserve's original 2% point-target for inflation was one in which bygones are bygones: That is, it would continue to try to achieve a 2% annual inflation rate no matter what had happened to inflation in the past. The Fed announcement in August 2020 that it would now target an *average* inflation rate of 2% meant that bygones were no longer bygones because past inflation would affect its target in the short run. If inflation had been running below the 2% target level, as it had prior to 2020, then average inflation would fall below 2%, and so to raise the average back to 2%, the Fed would seek to achieve an inflation rate above 2% for a short period of time. This would require the Fed to pursue easier monetary policy than it would have otherwise. If, on the other hand, inflation had been running above 2%, then the Fed would temporarily shoot for an inflation rate below 2% and pursue a tighter monetary policy.

There are two major advantages to this new monetary policy strategy. First, it would make it less likely that inflation expectations would drift down below the 2% level because past undershoots of the target would be made up over time by pursuing easier monetary policy. The second advantage of this new monetary policy strategy is that it would generate an automatic stabilizer for the economy. When a negative shock that caused inflation to fall below the 2% target occurred, average inflation targeting would commit the Fed to temporarily raise inflation to above 2%. Inflation expectations would then likely rise above the 2% level temporarily, thus lowering the real interest rate automatically even if the Fed did not or could not lower the federal funds rate.

There is one major objection to this new monetary policy strategy. If the Fed allowed inflation to temporarily rise above the 2% level, there might be concerns that the Fed is no longer committed to keeping inflation at the 2% level in the long run. To avoid this problem, the Fed would need to convince the public that an overshoot of the 2% inflation objective would not weaken the Fed's commitment to stabilizing inflation at the 2% level in the long run.

17.6 SHOULD CENTRAL BANKS TRY TO STOP ASSET-PRICE BUBBLES?

LO 17.6 Summarize the arguments for and against central bank policy responses to asset-price bubbles.

Over the centuries, economies have been subject to periodic **asset-price bubbles**, pronounced increases in asset prices, or "bubbles," that depart from fundamental values and that eventually burst resoundingly. The story of the global financial crisis, discussed in Chapter 11, indicates just how costly these bubbles can be. The bursting of the asset-price bubble in the U.S. housing market brought down the financial system, leading to an economic downturn, a rise in unemployment, and direct hardship for families who were forced to leave their homes after foreclosures.

The high economic cost of asset-price bubbles raises the following questions: What should central banks do about them? Should they use monetary policy to try to pop the bubbles? Are there regulatory measures they can take to rein in asset-price bubbles? To answer these questions, we need to ask whether there are different kinds of bubbles that require different types of responses.

Two Types of Asset-Price Bubbles

To think about central banks' response to asset-price bubbles, we first need to look at the different types of bubbles and how each might best be addressed. Asset-price bubbles are of two types: One type is driven by credit, and the second type is driven purely by overly optimistic expectations (dubbed "irrational exuberance" by Alan Greenspan).

Credit-Driven Bubbles

When a credit boom begins, an asset-price bubble may begin to form. Easier-to-get credit can be used to purchase particular assets and thereby raise their prices. The rise in asset values, in turn, encourages further lending for these assets, either because it increases the value of collateral, making it easier to borrow, or because it raises the value of capital at financial institutions, which gives them more capacity to lend. The lending for these assets can then further increase demand for them and hence raise their prices even more. This feedback loop—in which a credit boom drives up asset prices, which in turn fuels the credit boom, which drives asset prices even higher, and so on—can generate a bubble in which asset prices rise well above their fundamental values.

Credit-driven bubbles are particularly dangerous, as the recent global financial crisis has demonstrated. When asset prices come back down to earth and the bubble bursts, the collapse in asset prices leads to a reversal of the feedback loop: Loans go sour, lenders cut back on the credit supply, the demand for assets declines further, and prices drop even more. These were exactly the dynamics that characterized housing markets during the global financial crisis. Driven by a credit boom in subprime lending, housing prices rose way above fundamental values, but when housing prices crashed, credit shrivelled up and housing prices plummeted.

The resulting losses on subprime loans and securities eroded the balance sheets of financial institutions, causing a decline in credit (deleveraging) and a sharp fall in business and household spending and therefore in economic activity. During the global financial crisis, the interaction between housing prices and the health of financial institutions following the collapse of the housing price bubble endangered the operation of the financial system as a whole and had dire consequences for the economy.

Bubbles Driven Solely by Irrational Exuberance Bubbles that are driven solely by overly optimistic expectations, but that are not associated with a credit boom, pose much less risk to the financial system. For example, the bubble in technology stocks in the late 1990s was not fuelled by credit, and the bursting of the tech-stock bubble was not followed by a marked deterioration in financial institutions' balance sheets. The bursting of the tech-stock bubble thus did not have a very severe impact on the economy, and the recession that followed was quite mild. Bubbles driven solely by irrational exuberance are therefore far less dangerous than those driven by credit booms.

The Debate over Whether Central Banks Should Try to Pop Bubbles

Because asset prices are a central channel of monetary policy and directly affect its outcomes (the transmission mechanisms of monetary policy are discussed in Chapter 26), monetary policy certainly needs to respond to asset prices in order to obtain good outcomes in terms of inflation and output. Hence, the issue is not whether monetary policy should respond to asset-price movements at all but whether it should respond at a level over and above the level called for in terms of the objectives of stabilizing inflation and employment. Should monetary policy try to pop, or slow the growth of, potential asset-price bubbles to minimize damage to the economy when these bubbles burst? Alternatively, rather than responding directly to possible asset-price bubbles, should the monetary authorities respond only to the asset-price declines that occur after a bubble bursts, to stabilize both output and inflation? These opposing positions have been characterized as *leaning against* asset-price bubbles versus *cleaning up* after the bubbles burst, and so the debate over what to do about asset-price bubbles has been labelled the "lean versus clean" debate.

Whether central banks should try to pop, or prick, bubbles was actively debated before the global financial crisis, with Alan Greenspan arguing against such actions. Greenspan's position held great sway in central banking circles before the crisis. However, the crisis has led economists to reevaluate this viewpoint, and we look at the pro and con arguments below.

Con: Why Central Banks Should Not Try to Prick Asset-Price Bubbles But Should Just Clean up After They Burst Alan Greenspan's argument that central banks should not take actions to prick bubbles became known as the "Greenspan doctrine." His position reflected five arguments:

1. Asset-price bubbles are nearly impossible to identify. If central banks or government officials knew that a bubble was in progress, wouldn't market participants know as well? If so, then a bubble would be unlikely to develop, because market participants would know that prices were getting out of line with fundamentals. Unless central bank or government officials are smarter than market participants, an unlikely situation given the savvy of especially talented (and high-earning) market participants, they will be unlikely to identify when bubbles of this type are occurring. A strong argument, then, exists for not responding to suspected bubbles.
2. Although some economic analysis suggests that raising interest rates can diminish asset-price increases, raising interest rates may be very ineffective in restraining bubbles because market participants expect such high rates of return from buying bubble-driven assets. Furthermore, raising interest rates has often been found to cause a bubble to burst more severely, thereby increasing the damage to the economy. Another way of saying this is that bubbles are departures from normal behaviour, and it is unrealistic to expect that the usual tools of monetary policy will be effective in abnormal conditions.

3. Many different asset prices exist, and at any one time a bubble may be present in only a fraction of asset markets. Monetary policy actions are a very blunt instrument in such a case, as such actions would be likely to affect asset prices in general rather than the specific assets that are experiencing a bubble.
4. Monetary policy actions to prick bubbles can have harmful effects on the aggregate economy. If interest rates are raised significantly in an effort to curtail a bubble, the economy will slow, people will lose jobs, and inflation might fall below its desirable level. Indeed, as arguments 2 and 3 suggest, the rise in interest rates necessary to prick a bubble may be so high that it can be done only at great cost to workers and the economy. This is not to say that monetary policy should not respond to asset prices per se. The level of asset prices does affect aggregate demand (discussed in Chapter 26) and thus the evolution of the economy. Monetary policy should react to fluctuations in asset prices to the extent that they affect inflation and economic activity.
5. As long as policymakers respond in a timely fashion, by easing monetary policy aggressively after an asset bubble bursts, the harmful effects of a bursting bubble can be kept at a manageable level. Indeed, the Greenspan Fed acted in exactly this way after the stock market crash of 1987 and the bursting of the tech bubble in the stock market in 2000. Aggressive easing after the stock market bubbles burst in 1987 and 2000 was highly successful. The U.S. economy did not enter a recession after the stock market crash of 1987, and the recession that followed the tech bubble burst in 2000 was very mild.

Pro: Why Central Banks Should Try to Pop Bubbles The recent financial crisis clearly demonstrated that the bursting of credit-driven bubbles can be not only extremely costly but also very hard to clean up. Furthermore, credit-driven bubbles can occur even if price and output stability exist in the period leading up to them. Indeed, as we have seen, price and output stability might actually encourage credit-driven bubbles because they lead market participants to underestimate the amount of risk present in the economy. The global financial crisis has therefore provided a much stronger case for leaning against potential bubbles than for just cleaning up after they burst.

However, the distinction between the two types of bubbles, one of which (credit-driven) is much more costly to the economy than the other, suggests that the lean versus clean debate may have been miscast. Rather than leaning against potential asset-price bubbles, which would include both credit-driven and irrational exuberance–type bubbles, the case is much stronger for leaning against credit *booms*, which would involve leaning against credit-driven asset-price bubbles but not against asset-price bubbles driven by irrational exuberance. It is much easier to identify credit booms than asset-price bubbles. When asset-price bubbles are rising rapidly at the same time that credit is booming, the likelihood is greater that asset prices are deviating from fundamentals because laxer credit standards are driving asset prices upward. In this case, central bank or government officials are more likely to identify that a boom is in progress; this was indeed the case during the housing market bubble in the United States, when central banks and government officials were aware that lenders had weakened lending standards and that credit extension in the mortgage markets was rising at abnormally high rates.

The case for leaning against credit-driven bubbles seems strong, but what policies would be most effective in restraining them?

Macroprudential Policies First, it is important to recognize that the key principle to consider in designing effective policies to lean against credit booms is that such policies must curb excessive risk taking. Only when risk taking is excessive are credit booms likely to develop, and so it is natural to look to prudential regulatory measures

to constrain credit booms. Regulatory policy to affect what is happening in credit markets in the aggregate is referred to as **macroprudential regulation**, and it does seem to be the right tool for reining in credit-driven bubbles.

Financial regulation and supervision, either by central banks or by other government entities, with the usual elements of a well-functioning prudential regulatory and supervisory system, as described in Chapter 9, can prevent the excessive risk taking that can trigger a credit boom, which in turn will lead to an asset-price bubble. These elements include adequate disclosure and capital requirements, prompt corrective action, close monitoring of financial institutions' risk management procedures, and close supervision to enforce compliance with regulations. More generally, regulation should focus on preventing leverage cycles. As the global financial crisis demonstrated, the rise in asset prices that accompanied the credit boom resulted in higher capital buffers at financial institutions, supporting further lending in the context of unchanging capital requirements, which led to higher asset prices, and so on; in the bust, the value of the capital dropped precipitously, leading to a cut in lending. Capital requirements that are countercyclical, that is, adjusted upward during a boom and downward during a bust, might help eliminate the pernicious feedback loops that promote credit-driven bubbles.

A rapid rise in asset prices accompanied by a credit boom provides a signal that market failures or poor financial regulation and supervision might be causing a bubble to form. Central banks and other government regulators can then consider implementing policies to rein in credit growth directly or can implement measures to make sure credit standards are sufficiently high.

Monetary Policy The fact that the low-interest-rate policies of the Federal Reserve from 2002 to 2005 were followed by excessive risk taking suggests to many that overly easy monetary policy might promote financial instability, as was discussed in Chapter 11. Although it is far from clear that the Federal Reserve is primarily to blame for the housing bubble in the United States, research does suggest that low interest rates can encourage excessive risk taking, in what has been called the "risk-taking channel of monetary policy." Low interest rates may increase the incentives for asset managers in financial institutions to search for higher yields and hence increase risk taking. Low interest rates may also increase the demand for assets, raising their prices and leading to increased valuation of collateral, which in turn encourages lenders to lend to riskier borrowers.

The risk-taking channel of monetary policy suggests that monetary policy should be used to lean against credit booms. However, many of the Greenspan doctrine's objections to using monetary policy to prick booms are still valid, so wouldn't it be better to use macroprudential supervision to constrain credit booms, leaving monetary policy to focus on price and output stability?

This argument would be quite strong if macroprudential policies were able to do the job. However, there are doubts on this score. Prudential supervision is subject to more political pressure than monetary policy because it affects the bottom line of financial institutions more directly. Thus these institutions have greater incentives to lobby politicians to discourage macroprudential policies that would rein in credit booms, particularly during a credit boom, when they are making the most money. In addition, financial institutions are often very good at finding loopholes to avoid regulation, as we discovered in Chapter 10, and so macroprudential supervision may not be effective. The possibility that macroprudential policies may not be implemented sufficiently well to constrain credit booms suggests that monetary policy may have to be used instead.

An important lesson from the global financial crisis is that central banks and other regulators should not have a laissez-faire attitude and let credit-driven bubbles proceed without any reaction. Figuring out how to do this well, however, is indeed a daunting task.

17.7 TACTICS: CHOOSING THE POLICY INSTRUMENT

LO 17.7 Describe and assess the three criteria for choosing a policy instrument.

Now that we are familiar with the alternative strategies for monetary policy, let's look at how monetary policy is conducted on a day-to-day basis. Central banks directly control the tools of monetary policy—open market buyback operations, the bank rate, the interest rate on reserves, quantitative easing, and forward guidance—but knowing the tools and the strategies for implementing a monetary policy does not tell us whether that policy is easy or tight. To ascertain whether a policy is easy or tight, we can observe the **policy instrument** (also called an **operating instrument**), a variable that responds to the central bank's tools and indicates the stance (easy or tight) of monetary policy. A central bank like the Bank of Canada has at its disposal two basic types of policy instruments: reserve aggregates (total reserves, nonborrowed reserves, the monetary base, and the nonborrowed base) and interest rates (the overnight interest rate and other short-term interest rates). (Central banks in small countries can choose another policy instrument, the exchange rate, but we leave this topic to Chapter 18.) The policy instrument might be linked to an **intermediate target**, such as a monetary aggregate like M1+ or a long-term interest rate. Intermediate targets stand between the policy instrument and the goals of monetary policy (e.g., price stability, output growth); they are not as directly affected by the tools of monetary policy but might be more closely linked to the goals of monetary policy. As a study aid, Figure 17-2 shows a schematic of the linkages among the tools of monetary policy, policy instruments, intermediate targets, and the goals of monetary policy.

As an example, suppose the central bank's employment and inflation goals are consistent with a nominal GDP growth rate of 5%. The central bank might believe that the 5% nominal GDP growth rate will be achieved by a 4% growth rate for M1+ (an intermediate target), which will in turn be achieved by a growth rate of 3% for nonborrowed reserves (the policy instrument). Alternatively, the central bank might believe that the best way to achieve its objectives would be to set the overnight interest rate (a policy instrument) at, say, 4%. Can the central bank choose to target both the nonborrowed reserves and the overnight-rate policy instruments at the same time? The answer

Mini-lecture

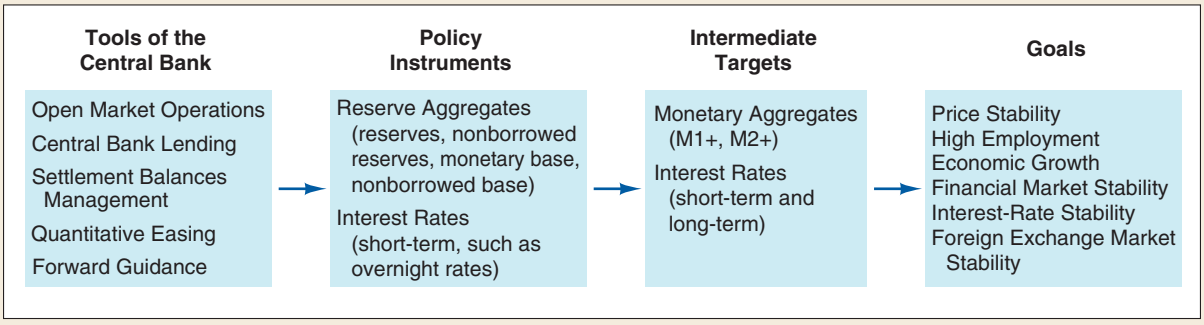


FIGURE 17-2 Linkages Among Central Bank Tools, Policy Instruments, Intermediate Targets, and Goals of Monetary Policy

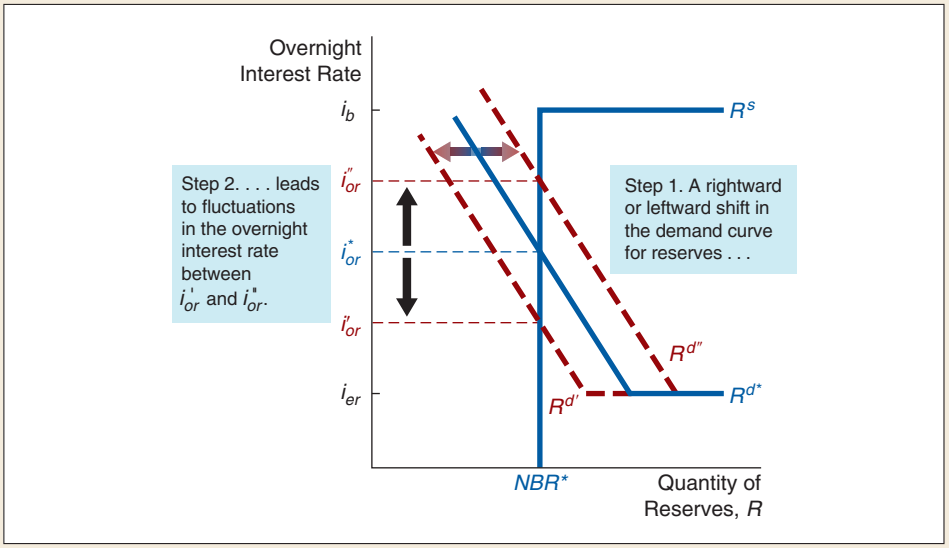
The tools of the central bank are used to change the policy instrument to achieve the intermediate target and then the goals of monetary policy.

Mini-lecture

FIGURE 17-3

Result of Targeting on Nonborrowed Reserves

Targeting on nonborrowed reserves of NBR^* will lead to fluctuations in the federal funds rate between i'_{ff} and i''_{ff} because of fluctuations in the demand for reserves between $R^{d'}$ and $R^{d''}$.



is no. The application of supply and demand analysis to the market for reserves we developed in Chapter 16 explains why a central bank must choose one or the other.

Let's first see why choosing an aggregate target involves losing control of the interest rate. Figure 17-3 contains a supply and demand diagram for the market for reserves. Although the central bank expects the demand curve for reserves to be at R^{d^*}, it fluctuates between $R^{d'}$ and $R^{d''}$ because of changes in banks' desire to hold excess reserves. If the central bank has a nonborrowed reserves target of NBR^* (say, because it has a target growth rate of the money supply of 4%), it expects that the overnight interest rate will be i^*_{or}. However, as the figure indicates, the fluctuations in the reserves demand curve between $R^{d'}$ and $R^{d''}$ will result in a fluctuation in the overnight interest rate between i'_{or} and i''_{or}. Pursuing an aggregate target implies that interest rates will fluctuate.

The supply and demand diagram in Figure 17-4 shows the consequences of an interest-rate target set at i^*_{or}. Again the central bank expects the reserves demand curve to be at R^{d^*}, but it fluctuates between $R^{d'}$ and $R^{d''}$ due to unexpected changes in deposits or banks' desire to hold reserves. If the demand curve rises to $R^{d''}$, the overnight interest rate will begin to rise above i^*_{or} and the central bank will engage in open market purchases of bonds until it raises the supply of nonborrowed reserves to NBR'', at which point the equilibrium overnight interest rate is again at i^*_{or}. Conversely, if the demand curve falls to $R^{d'}$ and lowers the overnight interest rate, the central bank would keep making open market sales until nonborrowed reserves fall to NBR'' and the overnight rate returns to i^*_{or}. The central bank's adherence to the interest-rate target thus leads to a fluctuating quantity of nonborrowed reserves and the money supply.

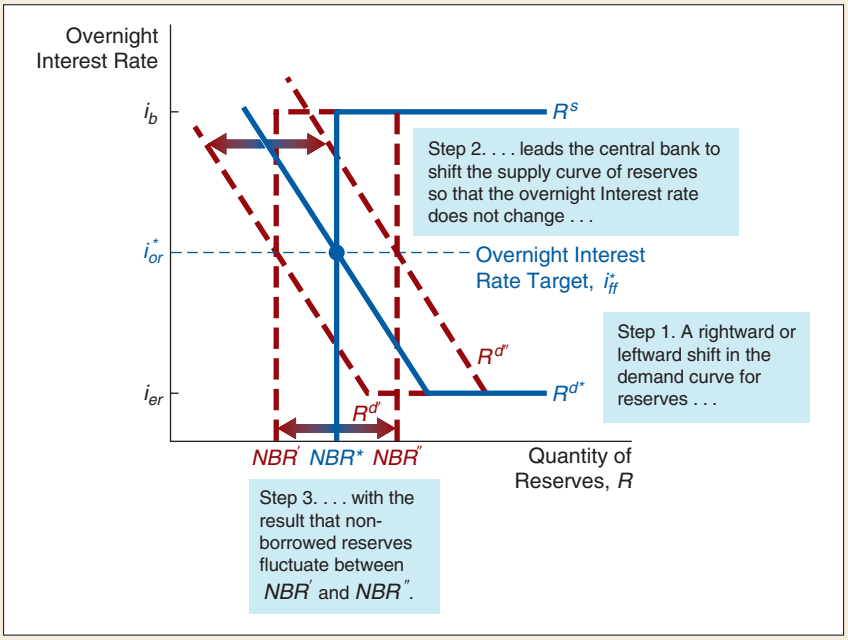
The conclusion from the supply and demand analysis is that interest-rate and reserve (monetary) aggregate targets are incompatible. A central bank can hit one or the other but not both. Because a choice between them has to be made, we need to examine what criteria should be used to select a policy instrument.

Mini-lecture

FIGURE 17-4

Result of Targeting on the Overnight Interest Rate

Targeting on the overnight interest rate i_{or}^{*} will lead to fluctuations in nonborrowed reserves between NBR' and NBR'' because of fluctuations in the demand for reserves between $R^{d'}$ and $R^{d''}$.


Overnight Interest Rate
i_b
i_{or}^{*}
i_{er}
R^s
Step 2. . . . leads the central bank to shift the supply curve of reserves so that the overnight Interest rate does not change . . .
Overnight Interest Rate Target, i_{ff}^{*}
$R^{d''}$
R^{d^*}
$R^{d'}$
Step 1. A rightward or leftward shift in the demand curve for reserves . . .
NBR' NBR^* NBR''
Quantity of Reserves, R
Step 3. . . . with the result that non-borrowed reserves fluctuate between NBR' and NBR''.

Criteria for Choosing the Policy Instrument

Three criteria apply when choosing a policy instrument: The instrument must be observable and measurable, it must be controllable by the central bank, and it must have a predictable effect on the goals.

Observability and Measurability Quick observability and accurate measurement of a policy instrument are necessary because it will be useful only if it signals the policy stance rapidly. Reserve aggregates like nonborrowed reserves are straightforward to measure, but some lag in reporting of reserve aggregates (a delay of two weeks) exists. Short-term interest rates, like the overnight interest rate, by contrast, not only are easy to measure but also are observable immediately. Thus, it seems that interest rates are more observable and measurable than are reserves and, therefore, are a better policy instrument.

However, as we learned in Chapter 4, the interest rate that is easiest to measure and observe is the nominal interest rate. It is typically a poor measure of the real cost of borrowing, which indicates with more certainty what will happen to the real GDP. This real cost of borrowing is more accurately measured by the real interest rate—that is, the nominal interest rate adjusted for expected inflation $r = i - \pi^{e}$. Unfortunately, real interest rates are extremely difficult to measure, because we do not have a direct way to measure expected inflation. Given that both interest rates and aggregates have observability and measurability problems, it is not clear whether one should be preferred over the other as a policy instrument.

Controllability A central bank must be able to exercise effective control over a variable if it is to function as a useful policy instrument. If the central bank cannot control the policy instrument, knowing that it is off track does little good, because the central bank has no way of getting it back on track.

Because of shifts in and out of currency, even reserve aggregates, such as nonborrowed reserves, are not completely controllable. Conversely, the Bank of Canada can control short-term interest rates, such as the overnight interest rate, very tightly. It might appear, therefore, that short-term interest rates would dominate reserve aggregates on the controllability criterion. However, a central bank cannot set short-term real interest rates because it does not have control over expectations of inflation. Once again, a clear-cut case cannot be made that short-term interest rates are preferable to reserve aggregates as a policy instrument, or vice versa.

Predictable Effect on Goals The most important characteristic of a policy instrument is that it must have a predictable effect on a goal. If a central bank can accurately and quickly measure the price of tea in China and can completely control its price, what good will that do? The central bank cannot use the price of tea in China to affect unemployment or the price level in its country. Because the ability to affect goals is so critical to the usefulness of any policy instrument, the tightness of the link from reserve or monetary aggregates to goals (output, employment, and inflation) or, alternatively, from interest rates to these goals, has been the subject of much research and debate. In recent years, most central banks have concluded that the link between interest rates and goals such as inflation is tighter than the link between reserve aggregates and inflation. For this reason, central banks throughout the world now generally use short-term interest rates as their policy instrument.

17.8 TACTICS: THE TAYLOR RULE

LO 17.8 Interpret and assess the performance of the Taylor rule as a hypothetical policy instrument for setting the overnight interest rate.

As we have seen, the Bank of Canada and most other central banks currently conduct monetary policy by setting a target for short-term interest rates like the overnight interest rate. But how should this target be chosen?

John Taylor of Stanford University has come up with an answer, called the **Taylor rule**. The Taylor rule indicates that the overnight interest rate should be set equal to the inflation rate plus an "equilibrium" real overnight interest rate (the real overnight rate that is consistent with full employment in the long run) plus a weighted average of two gaps: (1) an inflation gap—current inflation minus a target rate, and (2) an output gap—the percentage deviation of real GDP from an estimate of its potential (natural rate) level.[7] This rule can be written as follows:

$$\text{Overnight interest rate target} = \text{inflation rate} + \text{equilibrium real overnight rate} + \frac{1}{2}(\text{inflation gap}) + \frac{1}{2}(\text{output gap})$$

Taylor has assumed that the equilibrium real overnight interest rate is 2% and that an appropriate target for inflation would also be 2%, with equal weights of ½ on the inflation and output gaps. For a numerical example of the Taylor rule, suppose the inflation rate is at 3%, leading to a positive inflation gap of $1\%(3\% - 2\%)$, and real GDP is 1% above its potential, resulting in a positive output gap of 1%. Then the

[7]The original formulation of the Taylor rule can be found in John B. Taylor, "Discretion Versus Policy Rules in Practice," *Carnegie-Rochester Conference Series on Public Policy*, 39 (1993): 195–214. However, a more intuitive discussion with a historical perspective can be found in John B. Taylor, "A Historical Analysis of Monetary Policy Rules," in *Monetary Policy Rules*, ed. John B. Taylor (Chicago: University of Chicago Press, 1999), 319–341.

Taylor rule suggests that the overnight interest rate should be set at 6(= 3% inflation + 2% equilibrium real overnight rate + ½[1% inflation gap] + ½[1% output gap]).

An important feature of the Taylor rule is that the coefficient on the inflation gap, ½, is positive. If the inflation rate rises by 1 percentage point, then the overnight interest target is raised by 1.5 percentage points, and so by more than a one-to-one ratio. In other words, a rise in inflation of 1 percentage point leads to a real overnight rate increase of ½ percentage point. The principle that the monetary authorities should raise nominal interest rates by more than the increase in the inflation rate has been named the **Taylor principle**, and it is critical to the success of monetary policy. Suppose that the Taylor principle is not followed and nominal rates rise by *less* than the rise in the inflation rate so that real interest rates *fall* when inflation rises. Serious instability then results because a rise in inflation leads to an effective easing of monetary policy, which then leads to even higher inflation in the future. Indeed, this was a feature of monetary policy in the 1970s that led to a loss of the nominal anchor and the era of the so-called Great Inflation, when inflation rates climbed to double-digit levels. Fortunately, since 1979, the Taylor principle has become a feature of monetary policy, with much happier outcomes on both the inflation and the aggregate output fronts.

Some economists take the view that the presence of an output gap in the Taylor rule indicates that the Bank of Canada should care not only about keeping inflation under control but also about minimizing business-cycle fluctuations of output around its potential level. Caring about both inflation and output fluctuations is consistent with many statements by Bank of Canada officials that controlling inflation and stabilizing real output are important concerns of the Bank of Canada.

An alternative interpretation of the presence of the output gap in the Taylor rule is that the output gap is an indicator of future inflation, as stipulated in the **Phillips curve theory**. The Phillips curve theory indicates that changes in inflation are influenced by the state of the economy relative to its productive capacity as well as by other factors.

This productive capacity can be measured by potential GDP, which is a function of the natural rate of unemployment, the rate of unemployment consistent with full employment. A related concept is the **nonaccelerating inflation rate of unemployment (NAIRU)**—the rate of unemployment at which there is no tendency for inflation to change. Simply put, the theory states that when the unemployment rate is above NAIRU with output below potential, inflation will come down, but if it is below NAIRU with output above potential, inflation will rise. Prior to 1995, the NAIRU was thought to reside around 8%. However, with the decline in unemployment to around the 4% level in the late 1990s, with no increase (and even a slight decrease) in inflation, some critics have questioned the value of the Phillips curve theory. Either they claim that it just doesn't work anymore, or, alternatively, they believe that great uncertainty exists about the value of NAIRU. The Phillips curve theory is now highly controversial, and critics question whether it should be used as a guide for the conduct of monetary policy.

As Figure 17-5 shows, the Taylor rule does a pretty good, but not perfect, job of describing the Bank of Canada's setting of the overnight interest rate. Does this mean that the Bank of Canada should fire all its economists and put a computer in charge that simply has to compute the Taylor rule setting for the overnight rate? This would certainly save taxpayers a lot of money. Although the Bank of Canada does not do this, it does make use of the Taylor rule in thinking about how to conduct monetary policy. (See the FYI box "The Bank of Canada's Use of the Taylor Rule.") Given the above discussion, it is no surprise that the Taylor rule does not explain all the movements in the overnight interest rate in Figure 17-5. For this reason, financial institutions hire Bank of Canada watchers, as described in the FYI box "Bank of Canada Watchers."

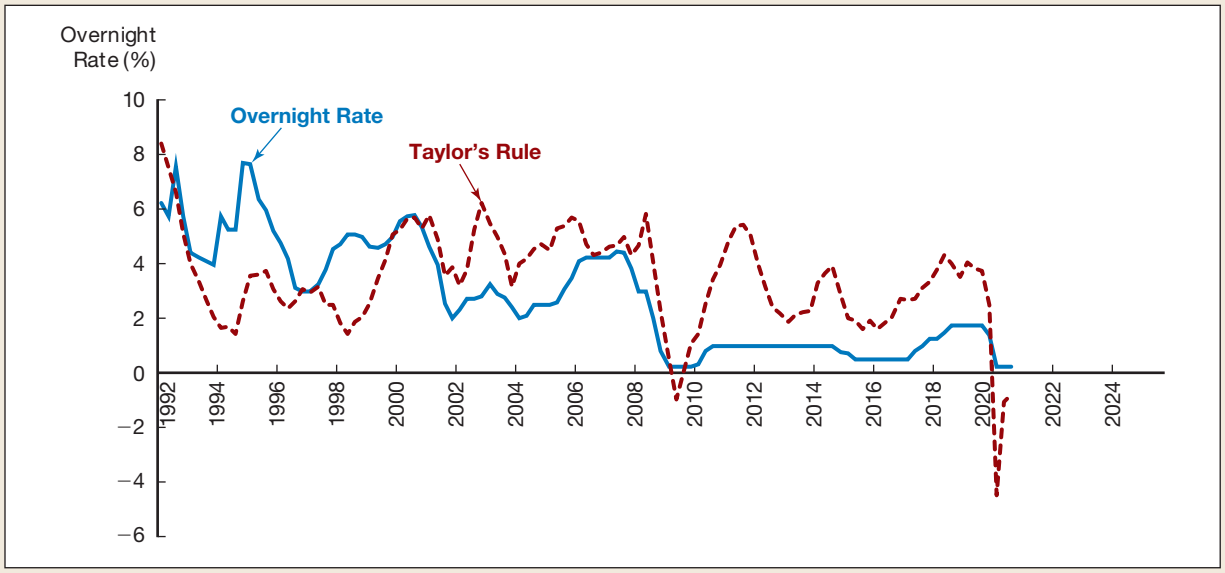


FIGURE 17-5 The Taylor Rule for the Overnight Interest Rate, 1992–2020

The Taylor rule does a pretty good job of describing the Bank of Canada's setting of the overnight interest rate.

Source: Statistics Canada CANSIM series V41690914 and V122514, Bank of Canada historical data, and authors' calculations.

FYI The Bank of Canada's Use of the Taylor Rule

Why hasn't the Bank of Canada put the overnight interest rate on the Taylor rule on autopilot, guided by a computer? There are several reasons why the Bank of Canada hasn't taken this drastic action. First and foremost, no perfect model of the economy exists, so even the best and brightest economists do not know the current inflation rate and output gap with certainty at any given moment. Since the economy is changing all the time, the Taylor rule coefficients are unlikely to stay constant in any event.

Even if we could determine these rates with certainty, monetary policy is by necessity a forward-looking activity because it takes a long time for monetary policy to affect the economy. Good monetary policy requires that the Bank of Canada forecast where inflation and economic activity are going to be in the future and then adjust the policy instrument accordingly. The Bank of Canada will therefore look at a much wider range of information than just the current inflation rate and output gap in setting policy, as is done in the Taylor rule. In other words, the conduct of monetary policy is as much an art as it is a science, requiring both careful analytics and human judgment. The Taylor rule leaves out all the art and so is unlikely to produce the best monetary policy outcomes. For example, financial crises, such as that which occurred from 2007 to 2009, require complex monetary policy actions because changes in credit spreads (the difference between interest rates on securities with credit risk and those without) may alter how the overnight interest rate affects investment decisions and therefore economic activity.

The bottom line is that putting monetary policy on autopilot with the Taylor rule with fixed coefficients would be a bad idea. The Taylor rule is, however, useful as a guide to monetary policy. If the setting of the policy instrument is very different from what the Taylor rule suggests, policymakers should ask whether they have a good reason for deviating from this rule. If they don't, then they might be making a mistake.

FYI Bank of Canada Watchers

As we have seen, the most important player in the determination of Canadian interest rates is the Bank of Canada. When the Bank of Canada wants to inject reserves into the system, it conducts open market purchases of bonds, which cause bond prices to increase and their interest rates to fall, at least in the short term. If the Bank withdraws reserves from the system, it sells bonds, thereby depressing their price and raising their interest rates. Knowing what actions the Bank of Canada might be taking can thus help investors and financial institutions to predict the future course of interest rates with greater accuracy. Because, as we have seen, changes in interest rates have a major impact on investors and financial institutions' profits, they are particularly interested in scrutinizing the Bank of Canada's behaviour. To assist in this task, financial institutions hire Bank of Canada watchers, experts on Bank of Canada behaviour who may have worked in the Bank of Canada and so have an insider's view of Bank of Canada operations. A Bank of Canada watcher who can accurately predict the course of monetary policy is a very valuable commodity, and successful Bank of Canada watchers therefore often earn very high salaries, well into the six-figure range and sometimes even higher.

SUMMARY

1. The six basic goals of monetary policy are price stability (the primary goal), high employment (output stability), economic growth, stability of financial markets, interest-rate stability, and stability in foreign exchange markets.
2. A strong nominal anchor is a key element of a successful monetary policy. A strong nominal anchor helps promote price stability by tying down inflation expectations and limiting the time-inconsistency problem, in which monetary policymakers conduct monetary policy in a discretionary way that focuses on short-run objectives but produces poor long-run outcomes.
3. Inflation targeting has several advantages: (1) By focusing the political debate on long-run inflation, it can reduce the likelihood of the time-inconsistency problem; (2) it is readily understood by the public and is highly transparent; (3) it increases accountability of the central bank; and (4) it results in more stable inflation. It does have some disadvantages, however: (1) Inflation is not easily controlled by the monetary authorities, and so an inflation target does not send immediate signals to the public and the markets; (2) it might impose a rigid rule on policymakers, although this has not been the case in practice; and (3) a sole focus on inflation may lead to larger output fluctuations, or lower economic growth, although this also has not been the case in practice.
4. Four lessons can be learned from the global financial crisis: (1) Developments in the financial sector have a far greater impact on economic activity than was earlier realized; (2) the effective lower bound on interest rates can be a serious problem; (3) the cost of cleaning up after a financial crisis is very high; and (4) price and output stability do not ensure financial stability.
5. The lessons of the global financial crisis provide support for more flexible inflation targeting, possibly with a higher inflation target.
6. The lessons from the crisis also suggest that monetary policy should lean against credit booms but not asset-price bubbles.
7. Because interest-rate and aggregate policy instruments are incompatible, a central bank must choose between them on the basis of three criteria: measurability, controllability, and their ability to affect goal variables predictably. Central banks now typically use short-term interest rates as their policy instrument.
8. The Taylor rule indicates that the overnight interest rate should be set equal to the inflation rate plus an "equilibrium" real overnight interest rate plus a weighted average of two gaps: (1) an inflation gap, current inflation minus a target rate; and (2) an output gap, the percentage deviation of real GDP from an estimate of its potential (natural rate) level. The output gap

in the Taylor rule can be interpreted as an indicator of future inflation, as stipulated in Phillips curve theory. However, this theory is controversial, because high output relative to potential as measured by low unemployment has not seemed to produce higher inflation in recent years.

MyLab Economics Make the grade with MyLab Economics: Select end-of-chapter questions can be found on MyLab Economics. You can practise them as often as you want, and most feature step-by-step guided instructions to help you find the right answer.

QUESTIONS

1. What are the benefits of using a nominal anchor for the conduct of monetary policy?
2. What incentives arise for a central bank to fall into the time-inconsistency trap of pursuing overly expansionary monetary policy?
3. Why would it be problematic for a central bank to have a primary goal of maximizing economic growth?
4. "Since financial crises can impart severe damage to the economy, a central bank's primary goal should be to ensure stability in financial markets." Is this statement true, false, or uncertain? Explain.
5. "A central bank with a dual mandate will achieve lower unemployment in the long run than a central bank with a hierarchical mandate in which price stability takes precedence." Is this statement true, false, or uncertain? Explain.
6. Why is a public announcement of numerical inflation rate objectives important to the success of an inflation-targeting central bank?
7. How does inflation targeting help reduce the time-inconsistency problem of discretionary policy?
8. What methods have inflation-targeting central banks used to increase communication with the public and to increase the transparency of monetary policymaking?
9. Why might inflation targeting increase support for the independence of the central bank in conducting monetary policy?
10. "Because inflation targeting focuses on achieving the inflation target, it will lead to excessive output fluctuations." Is this statement true, false, or uncertain? Explain.
11. What are the key advantages and disadvantages of the monetary strategy recently used by the Federal Reserve, in which the nominal anchor is implicit rather than explicit?
12. "The effective lower bound on short-term interest rates is not a problem, since the central bank can just use quantitative easing to lower intermediate and longer-term interest rates instead." Is this statement true, false, or uncertain? Explain.
13. If higher inflation is bad, then why might it be advantageous to have a higher inflation target rather than a lower target that is closer to zero?
14. Why might macroprudential regulation be more effective in managing asset-price bubbles than monetary policy?
15. Why might it be better to *lean* against credit-driven bubbles rather than just *clean* up after asset bubbles burst?
16. According to the Greenspan doctrine, under what conditions might a central bank respond to a perceived stock market bubble?
17. Classify each of the following as either a policy instrument or an intermediate target, and explain your choice.
    a. The ten-year Canada bond rate
    b. The monetary base
    c. M1+
    d. The overnight interest rate
18. "If the demand for reserves did not fluctuate, the Bank of Canada could pursue both a reserves target and an interest-rate target at the same time." Is this statement true, false, or uncertain? Explain.

19. What procedures can the Bank of Canada use to control the overnight interest rate? Why does control of this interest rate imply that the Bank of Canada will lose control of nonborrowed reserves?

20. Compare the monetary base to M1+ on the grounds of controllability and measurability. Which do you prefer as an intermediate target? Why?

21. "Interest rates can be measured more accurately and quickly than reserve aggregates; hence an interest rate is preferred to the reserve aggregates as a policy instrument." Do you agree or disagree? Explain your answer.

22. How can forward guidance as a tool of the central bank affect the policy instrument, intermediate targets, and goals?

23. What does the Taylor rule imply that policymakers should do to the overnight interest rate under the following scenarios?

a. Unemployment rises due to a recession.

b. An oil price shock causes the inflation rate to rise by 1% and output to fall by 1%.

c. The economy experiences prolonged increases in productivity growth while actual output growth is unchanged.

d. Potential output declines while actual output remains unchanged.

e. The Bank of Canada revises its inflation target downward.

f. The equilibrium real overnight interest rate decreases.

APPLIED PROBLEMS

24. If the Bank of Canada has an interest-rate target, why will an increase in the demand for reserves lead to a rise in the money supply? Use a graph of the market for reserves to explain.

25. Since monetary policy changes made through the overnight interest rate occur with a lag, policymakers are usually more concerned with adjusting policy according to changes in the forecasted or expected inflation rate rather than the current inflation rate. In light of this, suppose that monetary policymakers employ the Taylor rule to set the overnight interest rate, where the inflation gap is defined as the difference between expected inflation and the target inflation rate. Assume that the weights on both the inflation and output gaps are ½, the equilibrium real overnight interest rate is 2%, the inflation rate target is 2%, and the output gap is 1%.

a. If the expected inflation rate is 4%, then at what target should the overnight interest rate be set according to the Taylor rule?

b. Suppose half of Bank of Canada economists forecast inflation to be 3%, and half of Bank of Canada economists forecast inflation to be 5%. If the Bank of Canada uses the average of these two forecasts as its measure of expected inflation, then at what target should the overnight interest rate be set according to the Taylor rule?

c. Now suppose half of Bank of Canada economists forecast inflation to be 0%, and half forecast inflation to be 8%. If the Bank of Canada uses the average of these two forecasts as its measure of expected inflation, then at what target should the overnight interest rate be set according to the Taylor rule?

d. Given your answers to parts (a)–(c) above, do you think it is a good idea for monetary policymakers to use a strict interpretation of the Taylor rule as a basis for setting policy? Why or why not?

DATA ANALYSIS PROBLEMS

1. Go to the Bank of Canada's website and find inflation data (under "Statistics," then "Key Monetary Policy Variables") based on total CPI and core CPI-trim.

a. What was the 12-month growth rate for each of total CPI and core CPI-trim in December 2020?

b. How do the inflation rates in (a) compare to the inflation target?

2. Go to the Bank of Canada's website and find quarterly data (under "Indicators of Capacity and Inflation Pressures for Canada") on the output gap and the unemployment rate.

a. For the fourth quarter of 2020, is the output gap positive or negative?

b. Is the unemployment rate increasing or decreasing over the 2019 to 2020 period?

3. **Real-Time Data Analysis** Go to the St. Louis Federal Reserve FRED database and find data on the personal consumption expenditure price index (PCECTPI), real GDP (GDPC1), an estimate of potential GDP (GDPPOT), and the federal funds rate (FEDFUNDS). For the price index, adjust the *units* setting to "Percent Change From Year Ago" to convert the data to the inflation rate; for the federal funds rate, change the *frequency* setting to "Quarterly." Download the data into a spreadsheet. Assuming the inflation target is 2% and the equilibrium real fed funds rate is 2%, calculate the inflation gap and the output gap for each quarter, from 2000 until the most recent quarter of data available. Calculate the output gap as the percentage deviation of output from the potential level of output.
    a. Use the output and inflation gaps to calculate, for each quarter, the fed funds rate predicted by the Taylor rule. Assume that the weights on inflation stabilization and output stabilization are both ½ (see the formula in the chapter). Compare the current (quarterly average) federal funds rate to the federal funds rate prescribed by the Taylor rule. Does the Taylor rule accurately predict the current rate? Briefly comment.
    b. Create a graph that compares the predicted Taylor rule values with the actual quarterly federal-funds-rate averages. How well, in general, does the Taylor rule prediction fit the average federal funds rate? Briefly explain.
    c. Based on the results from the 2008–2009 period, explain the limitations of the Taylor rule as a formal policy tool. How do these limitations help explain the use of nonconventional monetary policy during this period?
    d. Suppose Congress changes the Fed's mandate to a hierarchical one in which inflation stabilization takes priority over output stabilization. In this context, recalculate the predicted Taylor rule value for each quarter since 2000, assuming that the weight on inflation stabilization is ¾ and the weight on output stabilization is ¼. Create a graph showing the Taylor rule prediction calculated in part (a), the prediction using the new "hierarchical" Taylor rule, and the fed funds rate. How, if at all, does changing the mandate change the predicted policy paths? How would the fed funds rate be affected by a hierarchical mandate? Briefly explain.
    e. Assume again equal weights of ½ on inflation and output stabilization, and suppose instead that beginning after the end of 2008, the equilibrium real fed funds rate declines by 0.05 each quarter (i.e., 2009:Q1 is 1.95, then 1.90, etc.), and once it reaches zero, it remains at zero thereafter. How does it affect the prescribed fed funds rate? Why might this be important for policymakers to take into consideration?

PART

6

International Finance and Monetary Policy

Crisis and Response: Foreign Exchange Market Turmoil and the IMF

From 2002 until 2008, the Canadian dollar steadily increased in value relative to the U.S. dollar. With the credit markets seizing up in September and October 2008, after the failure of Lehman Brothers, the U.S. dollar, instead of continuing its decline, appreciated sharply. The same "flight to quality" that led investors to step up their purchases of U.S. Treasury securities also led them to want to hold more U.S. dollars, thereby bidding up the U.S. dollar's value.

The good news for the U.S. dollar was often bad news for other currencies. Many countries in Latin America and Eastern Europe now found their currencies in free fall. The International Monetary Fund (IMF) stepped in and set up a new lending facility to make loans to distressed countries with fewer strings attached than were attached to the IMF's previous lending programs. The IMF started making loans to the tune of billions of dollars. The IMF, which had looked as though it was on the sidelines as the global financial crisis spread worldwide, now was moving to front and center.

The global financial crisis demonstrated that financial events that start in one country can have worldwide ramifications and that international financial institutions like the IMF have an important role in responding to such events, to make sure that the international financial system continues to work well. Chapter 18 outlines how the foreign exchange market functions and how exchange rates between different countries' currencies are determined. In Chapter 19, we examine how the international financial system operates and how it affects monetary policy.

18 The Foreign Exchange Market

Learning Objectives

18.1 Explain how the foreign exchange market works and why exchange rates are important.

18.2 Identify the main factors that affect exchange rates in the long run.

18.3 Draw the demand and supply curves for the foreign exchange market, and interpret the equilibrium in the market for foreign exchange.

18.4 List and illustrate the factors that affect the exchange rate in the short run.

Preview

On June 23, 2016, the United Kingdom voted to leave the European Union. This decision is referred to as *Brexit* (**Br**itian **Exit**). Within a day the British pound fell in value by 9% relative to the U.S. dollar.

The price of one currency in terms of another is called the *exchange rate*. As the Brexit example suggests, exchange rates are highly volatile. The exchange rate affects the economy and our daily lives because when the Canadian dollar becomes more valuable relative to foreign currencies, foreign goods become cheaper for Canadians, and Canadian goods become more expensive for foreigners. When the Canadian dollar falls in value, foreign goods become more expensive for Canadians, and Canadian goods become cheaper for foreigners.

Fluctuations in the exchange rate also affect both inflation and output and are an important concern for monetary policymakers. When the Canadian dollar falls in value, the higher prices of imported goods feed directly into a higher price level and inflation. At the same time, a declining Canadian dollar, which makes Canadian goods cheaper for foreigners, increases the demand for Canadian goods and leads to higher production and output.

How are currencies traded? What drives fluctuations in exchange rates? Why are exchange rates so volatile? We answer these questions by first examining the financial market in which currencies are traded. Next, we look at what affects exchange rates in the long run. We then develop a supply and demand analysis to explain what determines exchange rates in the short run. Finally, we use this supply and demand analysis to explain fluctuations in the exchange rate resulting from events such as Brexit and the global financial crisis.

18.1 FOREIGN EXCHANGE MARKET

LO 18.1 Explain how the foreign exchange market works and why exchange rates are important.

Most countries of the world have their own currencies: Canada has its dollar; the European Monetary Union, its euro; Brazil, its real; and China, its yuan. Trade between countries involves the mutual exchange of different currencies (or, more typically, bank deposits denominated in different currencies). When a Canadian firm buys foreign goods, services, or financial assets, for example, Canadian dollars (typically, bank deposits denominated in Canadian dollars) must be exchanged for foreign currency (bank deposits denominated in the foreign currency).

Following the Financial News — Foreign Exchange Rates

Foreign exchange rates are published daily in newspapers and on Internet sites such as www.finance.yahoo.com. Exchange rates for a currency such as the euro are quoted in two ways: in Canadian dollars per unit of domestic currency and in domestic currency per Canadian dollar. For example, on February 8, 2021, the euro exchange rate was quoted as $1.54 per euro and 0.64 euro per Canadian dollar. Canadians generally would regard the exchange rate with the euro as $1.54 per euro, while Europeans would think of it as 0.64 euro per dollar.

Exchange rates typically are quoted for the spot transaction (the spot exchange rate) and for forward transactions (the forward exchange rates) that will take place one month, three months, and six months in the future.

The trading of currencies and bank deposits denominated in particular currencies takes place in the **foreign exchange market**. Transactions conducted in the foreign exchange market determine the rates at which currencies are exchanged, which in turn determine the cost of purchasing foreign goods and financial assets.

What Are Foreign Exchange Rates?

An **exchange rate** is the price of one currency in terms of another. There are two kinds of exchange rate transactions. The predominant ones, called **spot transactions**, involve the immediate (two-day) exchange of bank deposits. **Forward transactions** involve the exchange of bank deposits at some specified future date. The **spot exchange rate** is the exchange rate for the spot transaction, and the **forward exchange rate** is the exchange rate for the forward transaction.

When a currency increases in value, it experiences **appreciation**; when it falls in value and is worth fewer Canadian dollars, it undergoes **depreciation**. At the beginning of 1999, for example, the euro was valued at 1.76 Canadian dollars; on February 8, 2021, it was valued at 1.54 dollars, as indicated in the Following the Financial News box "Foreign Exchange Rates." The euro *depreciated* by 12.5%: $(1.54 - 1.76)/1.76 = -0.125 = -12.5\%$. Equivalently, we could say that the Canadian dollar, which went from a value of 0.57 euro per dollar at the beginning of 1999 to a value of 0.64 euro per dollar on February 8, 2021, *appreciated* by 12.28%: $(0.64 - 0.57)/0.57 = 0.1228 = 12.28\%$.

Why Are Exchange Rates Important?

Exchange rates are important because they affect the relative prices of domestic and foreign goods. The dollar price of French goods to a Canadian is determined by the interaction of two factors: the price of French goods in euros and the euro/dollar exchange rate.

Suppose that Wanda the Winetaster, a Canadian, decides to buy a bottle of 1961 (a very good year) Château Lafite Rothschild to complete her wine cellar. If the price of the wine in France is 1000 euros and the exchange rate is $1.50 to the euro, the wine will cost Wanda $1500 (= 1000 euros × $1.50/euro). Now suppose that Wanda delays her purchase by two months, at which time the euro has appreciated to

per euro. If the domestic price of the bottle of Lafite Rothschild remains 1000 euros, its dollar cost will have risen from $1500 to $1600.

The same currency appreciation, however, makes the price of foreign goods in France less expensive. At an exchange rate of $1.50 per euro, a Dell computer priced at $2000 costs Pierre the Programmer 1333 euros; if the exchange rate increases to $1.60 per euro, the computer will cost only 1250 euros.

A depreciation of the euro lowers the cost of French goods in Canada but raises the cost of Canadian goods in France. If the euro drops in value to $1.00, Wanda's bottle of Lafite Rothschild will cost her only $1000 instead of $1500, and the Dell computer will cost Pierre 2000 euros rather than 1333 euros.

Such reasoning leads to the following conclusion: ***When a country's currency appreciates (rises in value relative to other currencies), the country's goods abroad become more expensive, and foreign goods in that country become cheaper (holding domestic prices constant in the two countries). Conversely, when a country's currency depreciates, its goods abroad become cheaper, and foreign goods in that country become more expensive.***

Depreciation of a currency makes it easier for domestic manufacturers to sell their goods abroad and makes foreign goods less competitive in domestic markets. The depreciation of the Canadian dollar from 2012 to 2016 helped Canadian industries sell more goods, but it hurt Canadian consumers because foreign goods were more expensive. The prices of French wine and cheese and the cost of vacationing abroad all rose as a result of the weak Canadian dollar.

In contrast, the appreciation of the Canadian dollar in 2021 made goods and services produced in Canada less competitive. However, the stronger dollar was positive for Canadian consumers because foreign goods like French cheese and vacationing abroad were less expensive.

How Is Foreign Exchange Traded?

You cannot go to a centralized location to watch exchange rates being determined; currencies are not traded on exchanges such as the Toronto Stock Exchange or the New York Stock Exchange. Instead, the foreign exchange market is organized as an over-the-counter market in which several hundred dealers (mostly banks) stand ready to buy and sell deposits denominated in foreign currencies. Because these dealers are in constant telephone and computer contact, the market is very competitive; in effect, it functions no differently from a centralized market.

An important point to note is that although banks, companies, and governments talk about buying and selling currencies in foreign exchange markets, they do not take a fistful of dollar bills and sell them for British pound notes. Rather, most trades involve the buying and selling of bank deposits denominated in different currencies. So when we say that a bank is buying dollars in the foreign exchange market, what we actually mean is that the bank is buying *deposits denominated in dollars*. The volume in this market is colossal, exceeding $6 trillion per day.

Trades in the foreign exchange market consist of transactions in excess of $1 million. The market that determines the exchange rates given in the Following the Financial News box is not the market in which we would buy foreign currency for a trip abroad. Instead, we buy foreign currency in the retail market, from dealers such as Thomas Cook, American Express, or from banks. Because retail prices are higher than wholesale prices, when we buy foreign exchange, we obtain fewer units of foreign currency per dollar—that is, we pay a higher price for foreign currency—than the exchange rates quoted in the newspaper.

18.2 EXCHANGE RATES IN THE LONG RUN

LO 18.2 Identify the main factors that affect exchange rates in the long run.

Like the price of any good or asset in a free market, exchange rates are determined by the interaction of supply and demand. To simplify our analysis of exchange rates in a free market, we divide it into two parts. First, we examine how exchange rates are determined in the long run; then we use our knowledge of the long-run determinants of exchange rates to help us understand how they are determined in the short run.

Theory of Purchasing Power Parity

One of the most prominent theories of how exchange rates are determined is the **theory of purchasing power parity (PPP)**. It states that the exchange rate between any two countries' currencies is such that a basket of goods and services, wherever it is produced, costs the same in both countries.

Suppose that you are buying a basket of goods and services in Canada that costs $100. The same basket costs 10 000 yen in Japan. The theory of purchasing power parity says that the exchange rate between the dollar and the yen is therefore 100 yen per dollar. At this exchange rate, the basket of goods and services in Canada would cost 10 000 yen (= $100 × 100 yen per dollar), the same cost as in Japan. Similarly, the basket of goods and services in Japan would cost $100 (10 000 yen × $0.01 per yen), the same as in Canada.

What is the intuition behind purchasing power parity? Suppose the basket of goods in Japan were shipped to Canada and there were no transportation costs or barriers to trade. It would sell for the same dollar amount in Canada as it does in Japan. Similarly, if the Canadian basket of goods were shipped to Japan, it would cost the same amount in yen. Only an exchange rate of 100 yen per dollar ensures that the prices are consistent across the two countries.

Another way of thinking about purchasing power parity is through a concept called the **real exchange rate**, the rate at which domestic goods can be exchanged for foreign goods. In effect, the real exchange rate is the price of domestic goods relative to the price of foreign goods denominated in the domestic currency. For example, if a basket of goods in Toronto costs $50 while the same basket of goods in Tokyo costs $75 (because the basket of goods costs 7500 yen and the exchange rate is 100 yen per dollar), then the real exchange rate is 0.66 (= $50/$75). In our example, the real exchange rate is below 1.0, indicating that it is cheaper to buy the basket of goods in Canada than in Japan. The real exchange rate indicates whether a currency is relatively cheap or not. The theory of PPP can also be described in terms of the real exchange rate. PPP predicts that in the long run the real exchange rate is always equal to 1.0 so that the purchasing power of the dollar is the same as that of other currencies, such as the yen or the euro.

What happens if one country's price level changes relative to another's? For example, suppose that the price level in Japan rises by 10%, while the price level in Canada does not change. As a result, the price of the basket in Japan has risen by 10% relative to the Canadian price. Now the cost of the basket of goods and services in Japan would rise by 10% to 11 000 yen, while the price of the same basket in Canada would remain $100. For the Canadian and the Japanese baskets of goods to cost the same, the exchange rate would have to change to 110 yen per dollar. Then the

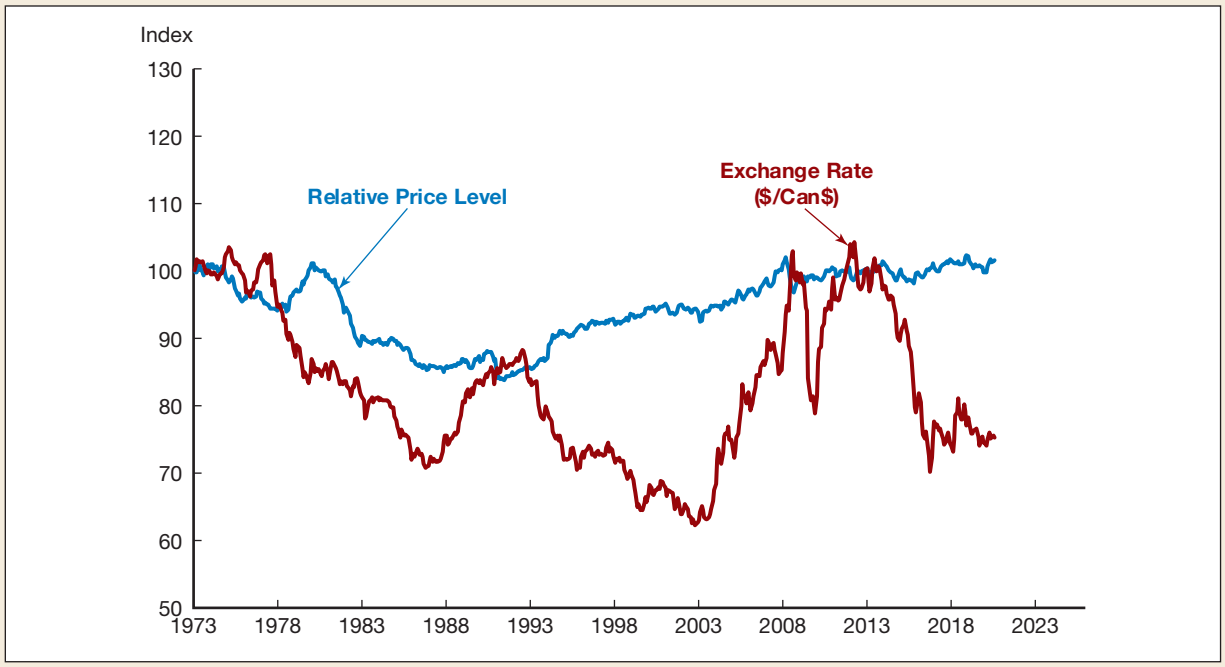


FIGURE 18-1 Purchasing Power Parity, Canada/United States, 1973–2020 (Index: March 1973 = 100)

The PPP relationship does not hold over short periods.

Source: Federal Reserve Bank of St. Louis FRED database.

Canadian basket would cost 11 000 yen in Japan ($100 × $100 yen per dollar), and the Japanese basket would cost $100 in Canada (11 000 yen × $0.009 per yen). Thus, according to the theory of PPP, if the price level in Japan rises 10% relative to the Canadian price level, the dollar appreciates by 10%.

Our Canadian/Japanese example demonstrates the key insight of the theory of PPP: ***If one country's price level rises relative to another's by a certain percentage, then the other country's currency appreciates by the same percentage.*** Because one currency's appreciation is the other currency's depreciation, there is an equivalent way of stating this result: ***If a country's price level rises relative to another's by a certain percentage, then its currency should depreciate by the same percentage.***

Evidence on Purchasing Power Parity As our Canadian/Japanese example demonstrates, the theory of PPP suggests that if one country's price level rises relative to another's, its currency should depreciate (and the other country's currency should appreciate). As you can see in Figure 18-1, this prediction is borne out only in the period before 1986 and in the 2000s. As the figure indicates, PPP theory has little predictive power in the short run. So even though PPP theory may provide some guidance

to the long-run movement of exchange rates, it is not perfect and in the short run is a particularly poor predictor.

What explains the PPP theory's failure to predict well in the short run?

Why the Theory of Purchasing Power Parity Cannot Fully Explain Exchange Rates There are three reasons why the theory of PPP does not fully explain exchange rates in the short run.

1. PPP theory does not take into account that many goods and services (whose prices are included in a measure of a country's price level) are **nontradable**; that is, they are not traded across borders. Housing, land, and services such as restaurant meals, haircuts, and golf lessons are not traded goods. So even though the prices of these items might rise, leading to a higher price level relative to another country's, the exchange rate would experience little direct effect.
2. Similar goods typically are not identical in both countries. For example, Toyota automobiles are not the same as Chevrolets, and so their prices do not have to be equal in a country. Toyotas can be more expensive relative to Chevys, and both Canadians and Japanese will still purchase Toyotas. Furthermore a rise in the price of Toyotas relative to Chevys will not necessarily mean that the yen must depreciate by the amount of the relative price increase of Toyotas over Chevys.
3. There are barriers to trade. One barrier to trade is transportation costs, but these costs have been declining dramatically over time. More important in today's world are government-imposed barriers to trade, such as **tariffs**, taxes on imported goods, and **quotas**, restrictions on the quantity of goods that can be imported. Imposition of a tariff or quota on an imported good can substantially raise its price relative to the price it would have elsewhere. For example, the U.S. government has imposed very high tariffs on imported sugar, resulting in U.S. sugar prices being several times higher than the price of sugar traded in other countries. Similarly, Japan has very high tariffs on imported rice, making it far more expensive in Japan than in other countries.

APPLICATION Burgernomics: Big Macs and PPP

Since 1986, *The Economist* magazine has published the Big Mac index as a "lighthearted guide to whether currencies are at their 'correct' level based on the theory of purchasing power parity." Big Macs are sold by McDonald's all around the world and are supposed to taste the same wherever they are sold. *The Economist* collects prices (in the local currency) of Big Macs sold in more than 50 different regions and countries. It then uses these prices to calculate both the exchange rate implied by PPP and the Big Mac index. Table 18-1 reproduces a portion of *The Economist*'s Big Mac numbers originally published in January 2020.

- Column (2) of Table 18-1 lists the local price of a Big Mac denominated in local currency.
- Column (3) lists the actual exchange rate with the U.S. dollar.

TABLE 18-1 Big Mac Index, January 2020

(1) Country	(2) Local Price of Big Mac	(3) Actual Exchange Rate (U.S. Dollars per Unit of Local Currency)	(4) Exchange Rate Implied by PPP (U.S. Dollars per Unit of Local Currency)	(5) Big Mac Index (Percent Difference Between Actual Exchange Rate and Implied PPP Exchange Rate)
Switzerland	6.50 Swiss francs	1.0328	0.8723	+18.4
Canada	6.77 Canadian dollars	0.7653	0.8375	−8.6
Euro area	4.12 euros	1.1121	1.3762	−19.2
Japan	390 yen	0.0091	0.0145	−37.5
China	21.5 yuan	0.1452	0.2637	−44.9
Russia	135 rubles	0.0162	0.0420	−61.2

Source: https://github.com/theeconomist/big-mac-data/releases/tag/2020-01.

- Column (4) provides the implied exchange rate if there is purchasing power parity (that is, if the local price of the Big Mac when converted to U.S. dollars equals the U.S. dollar price of the Big Mac in the United States, which was US$5.67 at the time).
- Column (5) lists the Big Mac index, the percent difference between the actual and implied exchange rates.

What conclusions can we draw from Table 18-1?

PPP Does Not Hold Exactly, but It Has Predictive Power Although the Big Mac prices indicate that PPP does not hold exactly, PPP does help predict exchange rates for the countries shown. According to PPP, when the Big Mac has a high price in terms of local currency, then the exchange rate quoted in U.S. dollars per unit of local currency should be low.

We see this relationship in the second and fourth columns of the table. Japan has the highest Big Mac price in terms of the local currency (390 yen), and it has the lowest exchange rate, US$0.0091 per yen. The euro area has the lowest Big Mac price in terms of the local currency (4.12 euros), and it has the highest exchange rate, US$1.1121 per euro.

Departures from PPP: Overvaluations and Undervaluations The exchange rate implied by PPP often departs substantially from the actual exchange rate. This should not be surprising given that Big Macs are nontradable. Think about taking a Big Mac from China to the United States. By the time you got it to the United States, it would be disgusting and might even give you food poisoning. The Big Mac index tells us how big the discrepancy is between the actual exchange rate and the exchange rate implied by PPP.

For example, the Big Mac index tells us that Switzerland has an actual exchange rate that is 18.4% higher than the exchange rate implied by PPP. As anyone who has visited Switzerland knows, this is a very expensive country to visit. The price of a Big Mac in Switzerland when converted into U.S. dollars using Switzerland's actual exchange rate is US$6.71, which is 18.4% higher than the US$5.67 you pay in the United States.

When a country's goods and services are expensive relative to other countries', we say that its currency is *overvalued* in terms of purchasing power parity.

Russia has the largest percentage difference between the actual exchange rate and the exchange rate implied by PPP. The Big Mac index indicates that the actual exchange rate is 61.2% lower than the exchange rate implied by PPP. In other words, if you are willing to fly to Russia to eat a Big Mac, you can get one for the incredibly low price of US$2.17.

When a country's goods and services are cheap relative to other countries', as in the case of Russia, we say that its currency (the Russian ruble) is *undervalued* in terms of purchasing power parity. ◆

Factors That Affect Exchange Rates in the Long Run

In the long run, four major factors affect the exchange rate: relative price levels, trade barriers, preferences for domestic versus foreign goods, and productivity. We examine how each of these factors affects the exchange rate while holding the other factors constant.

The basic reasoning proceeds along the following lines: Anything that increases the demand for domestically produced goods that are traded relative to foreign traded goods tends to appreciate the domestic currency, because domestic goods will continue to sell well even when the value of the domestic currency is higher. Similarly, anything that increases the demand for foreign goods relative to domestic goods tends to depreciate the domestic currency, because domestic goods will continue to sell well only if the value of the domestic currency is lower. In other words, ***if a factor increases the demand for domestic goods relative to foreign goods, the domestic currency will appreciate; if a factor decreases the relative demand for domestic goods, the domestic currency will depreciate.***

Relative Price Levels In line with PPP theory, when prices of Canadian goods rise (holding prices of foreign goods constant), the demand for Canadian goods falls, and the dollar tends to depreciate so that Canadian goods can still sell well. By contrast, if prices of Japanese goods rise, causing the relative prices of Canadian goods to fall, the demand for Canadian goods increases, and the dollar tends to appreciate because Canadian goods will continue to sell well even with a higher value of the domestic currency. ***In the long run, a rise in a country's price level (relative to the foreign price level) causes its currency to depreciate, and a fall in the country's relative price level causes its currency to appreciate.***

Trade Barriers Barriers to free trade such as tariffs and quotas can affect the exchange rate. Suppose Canada increases its tariff or puts a lower quota on Japanese steel. These increases in trade barriers increase the demand for Canadian steel, and the dollar tends to appreciate because Canadian steel will still sell well even with a higher

SUMMARY TABLE 18-1

Factors That Affect Exchange Rates in the Long Run

Factor	Change in Factor	Response of the Exchange Rate, E^*
Domestic price level†	↑	↓
Trade barriers†	↑	↑
Import demand	↑	↓
Export demand	↑	↑
Productivity†	↑	↑

*Units of foreign currency per Canadian dollar: ↑ indicates domestic currency appreciation; ↓, depreciation.

†Relative to other countries.

Note: Only increases (↑) in the factors are shown; the effects of decreases in the variables on the exchange rate are the opposite of those indicated in the "Response" column.

value of the dollar. ***Increasing trade barriers causes a country's currency to appreciate in the long run.***

Preferences for Domestic Versus Foreign Goods If the Japanese develop an appetite for Canadian goods—say, for Bombardier's high-speed trains and Canadian beef and pork—the increased demand for Canadian goods (exports) tends to appreciate the dollar because the Canadian goods will continue to sell well even at a higher value of the dollar. Likewise, if Canadians decide that they prefer Japanese cars to Canadian cars, the increased demand for Japanese goods (imports) tends to depreciate the dollar. ***Increased demand for a country's exports causes its currency to appreciate in the long run; conversely, increased demand for imports causes the domestic currency to depreciate.***

Productivity When productivity in a country rises, it tends to rise in domestic sectors that produce traded goods rather than nontraded goods. Higher productivity, therefore, is associated with a decline in the price of domestically produced traded goods relative to foreign traded goods. As a result, the demand for domestic traded goods rises, and the domestic currency tends to appreciate. If, however, a country's productivity lags behind that of other countries, its traded goods become relatively more expensive, and the currency tends to depreciate. ***In the long run, as a country becomes more productive relative to other countries, its currency appreciates.***[1]

Our long-run theory of exchange rate behaviour is summarized in Summary Table 18-1. We use the convention that the exchange rate E is quoted such that an appreciation of the domestic currency corresponds to a rise in the exchange rate. In the case of

[1]A country might be so small that a change in productivity or the preferences for domestic or foreign goods will have no effect on the prices of these goods relative to foreign goods. In this case, changes in productivity or changes in preferences for domestic or foreign goods affect the country's income but do not necessarily affect the value of the currency. In our analysis, we are assuming that changes in productivity or preferences can affect relative prices and consequently the exchange rate.

Canada, this means that we are quoting the exchange rate as units of foreign currency, say, euros, per Canadian dollar.[2]

18.3 EXCHANGE RATES IN THE SHORT RUN: A SUPPLY AND DEMAND ANALYSIS

LO 18.3 Draw the demand and supply curves for the foreign exchange market, and interpret the equilibrium in the market for foreign exchange.

We have developed a theory of the long-run behaviour of exchange rates. However, because factors driving long-run changes in exchange rates move slowly over time, if we are to understand why exchange rates exhibit such large changes (sometimes several percentage points) from day to day, we must develop a supply and demand analysis that explains how current exchange rates (spot exchange rates) are determined in the short run.

The key to understanding the short-run behaviour of exchange rates is to recognize that an exchange rate is the price of domestic assets (bank deposits, bonds, equities, and so on, denominated in the domestic currency) in terms of foreign assets (similar assets denominated in the foreign currency). Because the exchange rate is the price of one asset in terms of another, the natural way to investigate the short-run determination of exchange rates is with a supply and demand analysis that uses an asset market approach, which relies heavily on the theory of portfolio choice developed in Chapter 5. As you will see, however, the long-run determinants of the exchange rate we have just outlined also play an important part in the short-run asset market approach.

In the past, supply and demand approaches to exchange rate determination emphasized the role of import and export demand. The more modern asset market approach used here emphasizes stocks of assets rather than the flows of exports and imports over short periods because export and import transactions are small relative to the amounts of domestic and foreign assets held at any given time. For example, the total value of foreign exchange transactions in the United States each year is well over 25 times greater than the amount of U.S. exports and imports. Thus, over short periods, decisions to hold domestic or foreign assets have a much greater role in exchange rate determination than the demand for exports and imports does.

Supply Curve for Domestic Assets

We start by discussing the supply curve. In this analysis we treat Canada as the home country, so domestic assets are denominated in (Canadian) dollars. For simplicity, we use euros to stand for any foreign country's currency, so foreign assets are denominated in euros.

The quantity of dollar assets supplied is primarily the quantity of bank deposits, bonds, and equities in Canada, and for all practical purposes we can take this amount as fixed with respect to the exchange rate. The quantity supplied at any exchange rate is the same, so the supply curve, S, is vertical, as shown in Figure 18-2.

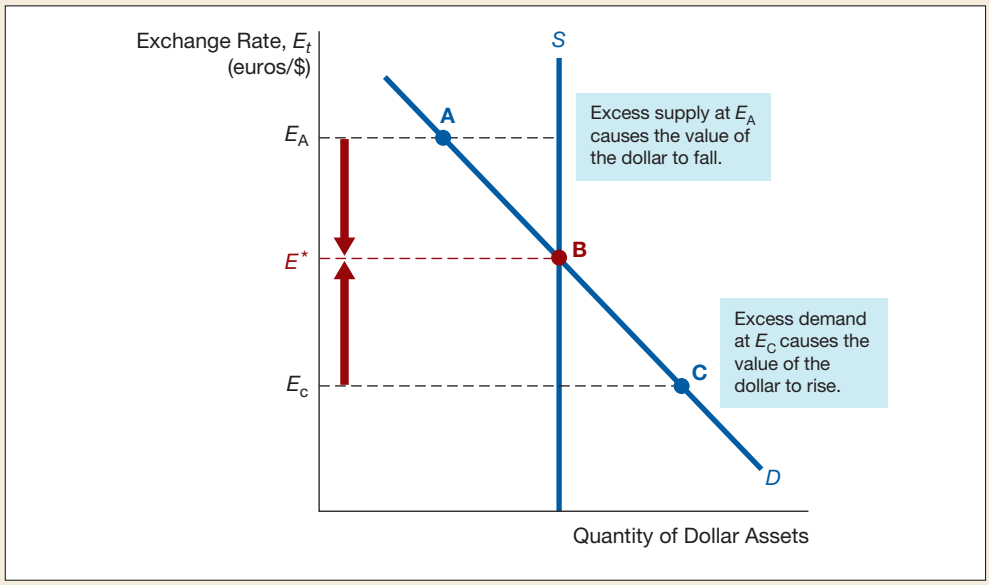
[2]Exchange rates can be quoted either as units of foreign currency per domestic currency or as units of domestic currency per foreign currency. In professional writing, many economists quote exchange rates as units of domestic currency per foreign currency so that an appreciation of the domestic currency is portrayed as a fall in the exchange rate. The opposite convention is used in this text because it is more intuitive to think of an appreciation of the domestic currency as a rise in the exchange rate.

Mini-lecture

FIGURE 18-2

Equilibrium in the Foreign Exchange Market

Equilibrium in the foreign exchange market occurs at point B, the intersection of the demand curve D and the supply curve S at an exchange rate of E^*.



Demand Curve for Domestic Assets

The demand curve traces out the quantity of domestic (dollar) assets demanded at each current exchange rate by holding everything else constant, particularly the expected future value of the exchange rate. We write the current exchange rate (the spot exchange rate) as E_t and the expected exchange rate for the next period as E^e_{t+1}. As suggested by the theory of portfolio choice, the most important determinant of the quantity of domestic (dollar) assets demanded is the relative expected return on domestic assets. Let's see what happens as the current exchange rate E_t falls.

Suppose we start at point A in Figure 18-2, where the current exchange rate is at E_A. With the future expected value of the exchange rate held constant at E^e_{t+1}, a lower value of the current exchange rate—say, at E^*—implies that the dollar is more likely to rise in value, that is, appreciate. The greater the expected rise (appreciation) of the dollar, the higher the relative expected return on dollar (domestic) assets. According to the theory of portfolio choice, because dollar assets are now more desirable to hold, the quantity of dollar assets demanded will rise, as shown by point B in Figure 18-2. If the current exchange rate falls even further to E_C, there will be an even higher expected appreciation of the dollar, a higher expected return, and therefore an even greater quantity of dollar assets demanded. This effect is shown at point C at Figure 18-2. The resulting demand curve D, which connects these points, is downward-sloping, indicating that at lower current values of the dollar (everything else being equal), the quantity demanded of dollar assets is higher.

A numerical example may make this clearer. Suppose that the future expected value of the dollar, E^e_{t+1}, is 1.20 euros per dollar, and E_A is 1.10 euros per dollar, E^* is 1.00 euros per dollar and E_C is 0.90 euro per dollar. At E_A where the exchange rate is 1.10 euros per dollar, the expected appreciation of the dollar is 10% $(= (1.20 - 1.10)/1.10 = +0.10 = +10\%)$. When the exchange rate drops to E^* of 1.00 euros per dollar, the expected appreciation is now a larger 20% $(= (1.20 - 1.00)/1.00 = +0.20 = +20\%)$, and so the quantity of dollar assets demanded is higher. When the exchange rate declines even further to E_C of 0.90 euro per dollar, the expected appreciation of the dollar

rises further to 30% $(= (1.20 - 0.90)/0.90 = +0.30 = +30\%)$ and so the quantity demanded of dollar assets is even higher.

Equilibrium in the Foreign Exchange Market

As in the usual supply and demand analysis, the market is in equilibrium when the quantity of dollar assets demanded equals the quantity supplied. In Figure 18-2, equilibrium occurs at point B, the intersection of the demand and supply curves. At point B, the exchange rate is E^*.

Suppose the exchange rate is at E_A, which is higher than the equilibrium exchange rate of E^*. As we can see in Figure 18-2, the quantity of dollar assets supplied is now greater than the quantity demanded, a condition of excess supply. Given that more people want to sell dollar assets than want to buy them, the value of the dollar will fall. As long as the exchange rate remains above the equilibrium exchange rate, an excess supply of dollar assets will continue to be available, and the dollar will fall in value until it reaches the equilibrium exchange rate of E^*.

Similarly, if the exchange rate is less than the equilibrium exchange rate at E_C, the quantity of dollar assets demanded will exceed the quantity supplied, a condition of excess demand. Given that more people want to buy dollar assets than want to sell them, the value of the dollar will rise until the excess demand disappears and the value of the dollar is again at the equilibrium exchange rate of E^*.

18.4 EXPLAINING CHANGES IN EXCHANGE RATES

LO 18.4 List and illustrate the factors that affect the exchange rate in the short run.

The supply and demand analysis of the foreign exchange market illustrates how and why exchange rates change.[3] We have simplified this analysis by assuming that the amount of dollar assets is fixed: The supply curve is vertical at a given quantity and does not shift. Under this assumption, we need to look at only those factors that shift the demand curve for dollar assets to explain how exchange rates change over time.

Shifts in the Demand for Domestic Assets

As we have seen, the quantity of domestic (dollar) assets demanded depends on the relative expected return on dollar assets. To see how the demand curve shifts, we need to determine how the quantity demanded changes, holding the current exchange rate, E_t, constant when other factors change.

For insight into the direction in which the demand curve will shift, suppose you are an investor who is considering putting funds into domestic (dollar) assets. When a factor changes, you must decide whether, at a given level of the current exchange rate and holding all other variables constant, you would earn a higher or lower expected return on dollar assets versus foreign assets. This decision will tell you whether you want to hold more or fewer dollar assets and thus whether the quantity demanded will increase or decrease at each level of the exchange rate. The direction of the change in

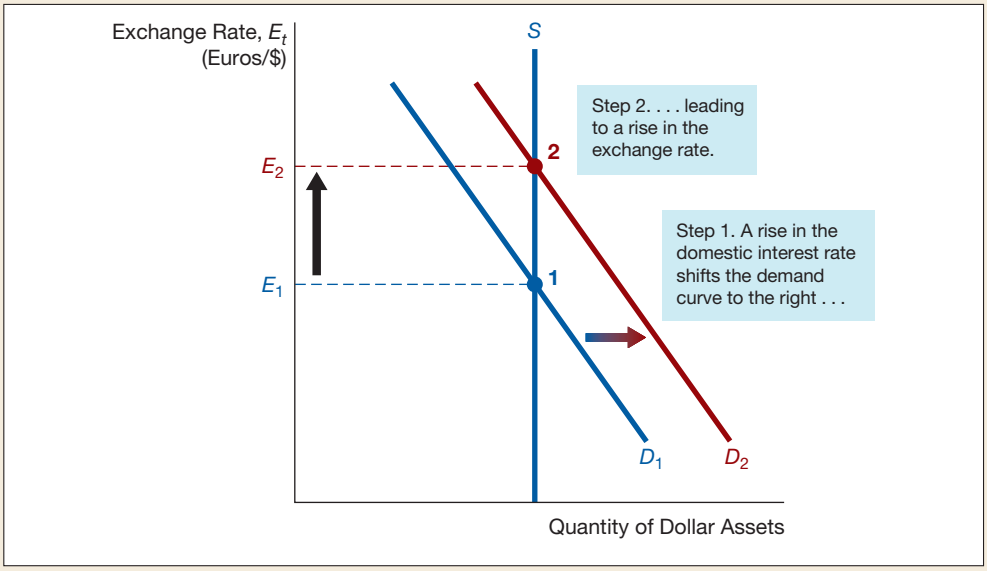
[3]How and why exchange rates change can also be modelled using the *interest parity condition*, an important concept in international finance that shows the relationships among domestic interest rates, foreign interest rates, and the expected appreciation of domestic currency. The interest parity condition and how it explains the determination of exchange rates is discussed in an appendix to this chapter found online.

Mini-lecture

FIGURE 18-3

Response to an Increase in the Domestic Interest Rate, i^D

When the domestic interest rate i^D increases, the relative expected return on domestic (dollar) assets increases and the demand curve shifts to the right. The equilibrium exchange rate rises from E_1 to E_2.


Exchange Rate, E_t (Euros/$)
S
Step 2. . . . leading to a rise in the exchange rate.
E_2
2
Step 1. A rise in the domestic interest rate shifts the demand curve to the right . . .
E_1
1
D_1
D_2
Quantity of Dollar Assets

the quantity demanded at each exchange rate indicates which way the demand curve will shift. In other words, if the relative expected return on dollar assets rises, holding the current exchange rate constant, the demand curve will shift to the right. If the relative expected return falls, the demand curve will shift to the left.

Domestic Interest Rate, i^D Suppose that dollar assets pay an interest rate of i^D. When the domestic interest rate on dollar assets i^D rises, holding the current exchange rate E_t and everything else constant, the return on dollar assets increases relative to the return on foreign assets, and so people will want to hold more dollar assets. The quantity of dollar assets demanded increases at every value of the exchange rate, as shown by the rightward shift of the demand curve from D_1 to D_2 in Figure 18-3. The new equilibrium is reached at point 2, the intersection of D_2 and S, and the equilibrium exchange rate rises from E_1 to E_2. ***An increase in the domestic interest rate i^D shifts the demand curve for domestic assets D to the right and causes the domestic currency to appreciate*** ($E\uparrow$).

Conversely, if i^D falls, the relative expected return on dollar assets falls, the demand curve shifts to the left, and the exchange rate falls. ***A decrease in the domestic interest rate i^D shifts the demand curve for domestic assets D to the left and causes the domestic currency to depreciate*** ($E\downarrow$).

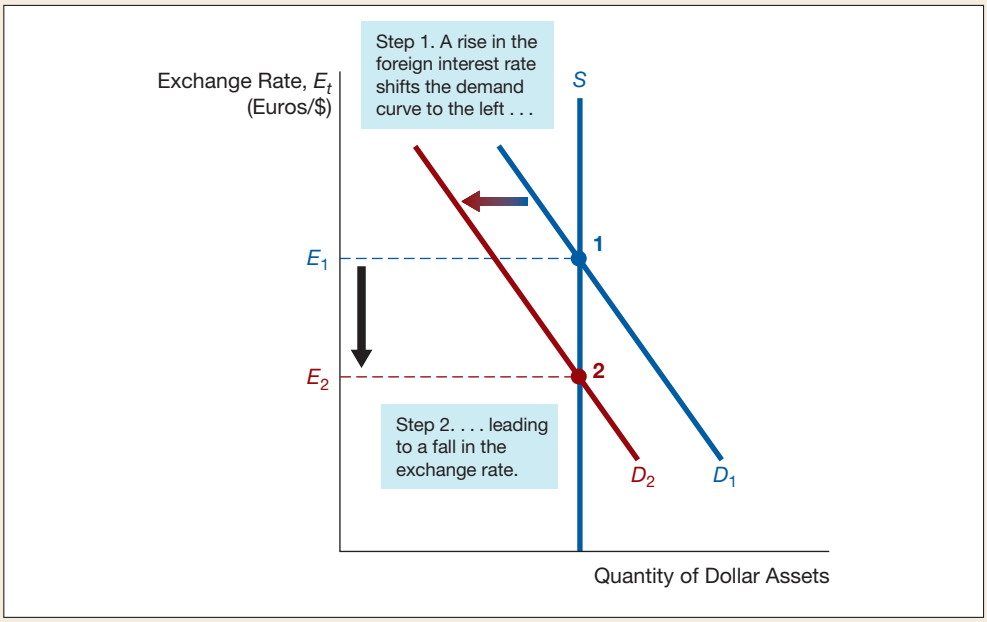
Foreign Interest Rate, i^F Suppose foreign assets pay an interest rate of i^F. When the foreign interest rate i^F rises, holding the current exchange rate and everything else constant, the return on foreign assets rises relative to the return on dollar assets. Thus the relative expected return on dollar assets falls. Now people want to hold fewer dollar assets, and the quantity demanded decreases at every value of the exchange rate. This scenario is shown by the leftward shift of the demand curve from D_1 to D_2 in Figure 18-4. The new equilibrium is reached at point 2, when the value of the dollar has fallen. Conversely, a decrease in i^F raises the relative expected return on dollar assets, shifts the demand curve to the right, and raises the exchange rate. To summarize, ***an increase in***

Mini-lecture

FIGURE 18-4

Response to an Increase in the Foreign Interest Rate, i^F

When the foreign interest rate i^F increases, the relative expected return on domestic (dollar) assets falls and the demand curve shifts to the left. The equilibrium exchange rate falls from E_1 to E_2.


Step 1. A rise in the foreign interest rate shifts the demand curve to the left . . .
Exchange Rate, E_t (Euros/$)
S
E_1
1
E_2
2
Step 2. . . . leading to a fall in the exchange rate.
D_2
D_1
Quantity of Dollar Assets

the foreign interest rate i^F shifts the demand curve D to the left and causes the domestic currency to depreciate; a fall in the foreign interest rate i^F shifts the demand curve D to the right and causes the domestic currency to appreciate.

Changes in the Expected Future Exchange Rate, E^e_{t+1} Expectations about the future value of the exchange rate play an important role in shifting the current demand curve because the demand for domestic assets, like that for any physical or financial asset, depends on the future resale price. Given the current exchange rate E_t, any factor that causes the expected future exchange rate E^e_{t+1} to rise increases the expected appreciation of the dollar. The result is a higher relative expected return on dollar assets, which increases the demand for dollar assets at every exchange rate, thereby shifting the demand curve to the right from D_1 to D_2 in Figure 18-5. The equilibrium exchange rate rises to point 2 at the intersection of the D_2 and S curves. ***A rise in the expected future exchange rate E^e_{t+1} shifts the demand curve to the right and causes an appreciation of the domestic currency.*** According to the same reasoning, ***a fall in the expected future exchange rate E^e_{t+1} shifts the demand curve to the left and causes a depreciation of the currency.***

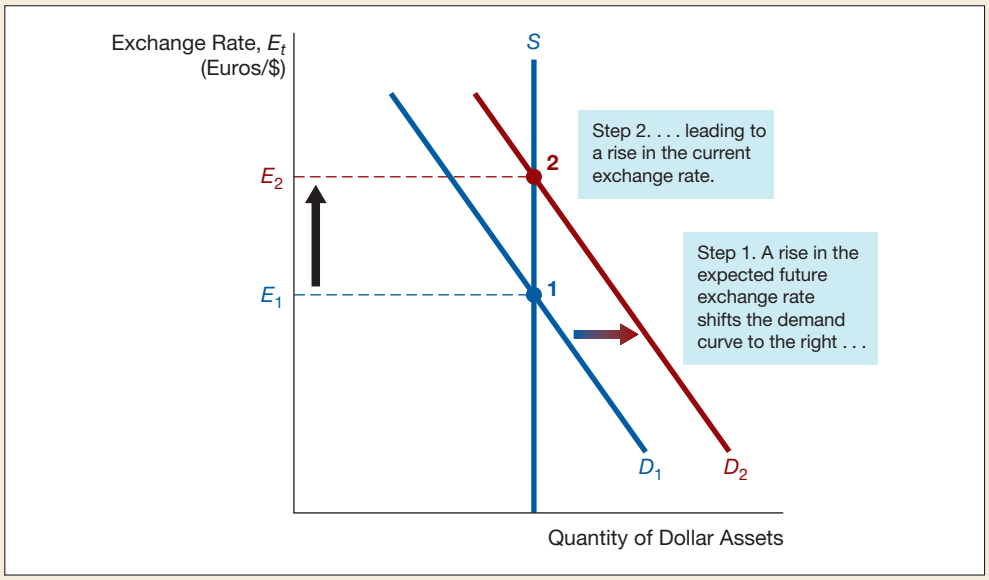
Earlier in the chapter we discussed the determinants of the exchange rate in the long run: the relative price level, relative trade barrier, import and export demand, and relative productivity (refer to Summary Table 18-1). These four factors influence the expected future exchange rate. The theory of purchasing power parity suggests that if a higher Canadian price level relative to the foreign price level is expected to persist, then the dollar will depreciate in the long run. A higher expected relative Canadian price level should thus have a tendency to lower E^e_{t+1}, lower the relative expected return on dollar assets, shift the demand curve to the left, and lower the current exchange rate.

Mini-lecture

FIGURE 18-5

Response to an Increase in the Expected Future Exchange Rate, E^e_{t+1}

When the expected future exchange rate increases, the relative expected return on domestic (dollar) assets rises and the demand curve shifts to the right. The equilibrium exchange rate rises from E_1 to E_2.


Exchange Rate, E_t (Euros/$)
S
Step 2. . . . leading to a rise in the current exchange rate.
2
E_2
Step 1. A rise in the expected future exchange rate shifts the demand curve to the right . . .
E_1
1
D_1
D_2
Quantity of Dollar Assets

Similarly, the other long-run determinants of the exchange rate can influence the relative expected return on dollar assets and the current exchange rate. Briefly, the following changes, all of which increase the demand for domestic goods relative to foreign goods, will raise E^e_{t+1}: (1) expectations of a fall in the Canadian price level relative to the foreign price level; (2) expectations of higher Canadian trade barriers relative to foreign trade barriers; (3) expectations of lower Canadian import demand; (4) expectations of higher foreign demand for Canadian exports; and (5) expectations of higher Canadian productivity relative to foreign productivity. By increasing E^e_{t+1}, all of these changes increase the relative expected return on dollar assets, shift the demand curve to the right, and cause an appreciation of the domestic currency, the dollar.

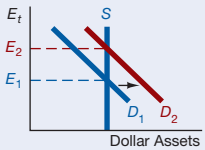
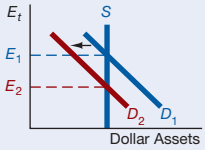
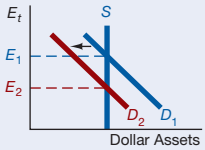
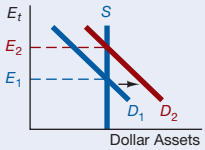
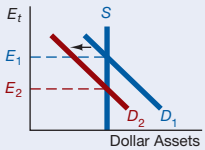
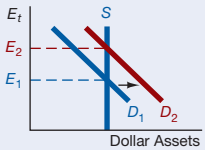
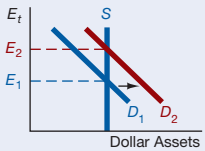
The fact that exchange rates are so volatile is explained by the analysis here. Because expected appreciation of the domestic currency affects the expected return on domestic assets, expectations about the price level, inflation, trade barriers, productivity, import demand, export demand, and monetary policy play important roles in determining the exchange rate. When expectations about any of these variables change, as they do—and often, at that—our analysis indicates that the expected return on domestic assets, and therefore the exchange rate, will be immediately affected. Because expectations on all these variables change with just about every bit of news that appears, it is not surprising that the exchange rate is volatile.

Recap: Factors That Change the Exchange Rate

Summary Table 18-2 outlines all of the factors that shift the demand curve for domestic assets and thereby cause the exchange rate to change. Shifts in the demand curve occur when one factor changes, holding everything else constant, including the current

SUMMARY TABLE 18-2

Factors That Shift the Demand Curve for Domestic Assets and Affect the Exchange Rate

Factor	Change in Factor	Change in Quantity Demanded of Domestic Assets at Each Exchange Rate	Response of Exchange Rate, E_t	
Domestic interest rate, i^D	↑	↑	↑	
Foreign interest rate, i^F	↑	↓	↓	
Expected domestic price level*	↑	↓	↓	
Expected trade barriers*	↑	↑	↑	
Expected import demand	↑	↓	↓	
Expected export demand	↑	↑	↑	
Expected productivity*	↑	↑	↑	

*Relative to other countries.

Note: Only increases (↑) in the factors are shown; the effects of decreases in the variables on the exchange rate are the opposite of those indicated in the "Response" column.

exchange rate. Again, the theory of portfolio choice asserts that changes in the relative expected return on dollar assets are the source of shifts in the demand curve.

Let's review what happens when each of the seven factors in Summary Table 18-2 changes. Remember that to understand the direction in which the demand curve will shift, we must consider what happens to the relative expected return on dollar assets when the factor changes. If the relative expected return rises, holding the current exchange rate constant, the demand curve shifts to the right. If the relative expected return falls, the demand curve shifts to the left.

1. When the interest rates on domestic assets i^D rise, the expected return on dollar assets rises at each exchange rate and so the quantity demanded increases. The demand curve therefore shifts to the right, and the equilibrium exchange rate rises, as shown in the first row of Summary Table 18-2.
2. When the foreign interest rate i^F rises, the return on foreign assets rises, and so the relative expected return on dollar assets falls. The quantity demanded of dollar assets then falls, the demand curve shifts to the left, and the exchange rate falls, as shown in the second row of Summary Table 18-2.
3. When the expected domestic price level rises relative to the foreign price level, our analysis of the long-run determinants of the exchange rate indicates that the value of the dollar will fall in the future. The expected return on dollar assets thus falls, the quantity demanded declines, the demand curve shifts to the left, and the exchange rate falls, as indicated in the third row of Summary Table 18-2.
4. With higher expected trade barriers, the value of the dollar is higher in the long run and the expected return on dollar assets is higher. The quantity demanded of dollar assets thus rises, the demand curve shifts to the right, and the exchange rate rises, as shown in the fourth row of Summary Table 18-2.
5. When the expected import demand rises, we expect the exchange rate to depreciate in the long run, so the expected return on dollar assets falls. The quantity demanded of dollar assets at each value of the current exchange rate therefore falls, the demand curve shifts to the left, and the exchange rate falls, as shown in the fifth row of Summary Table 18-2.
6. When the expected export demand rises, the exchange rate is expected to appreciate in the long run. Thus the expected return on dollar assets rises, the demand curve shifts to the right, and the exchange rate rises, as indicated in the sixth row of Summary Table 18-2.
7. When expected domestic productivity rises, the exchange rate is expected to appreciate in the long run, so the expected return on domestic assets rises. The quantity demanded at each exchange rate therefore rises, the demand curve shifts to the right, and the exchange rate rises, as shown in the seventh row of Summary Table 18-2.

APPLICATION Effects of Changes in Interest Rates on the Equilibrium Exchange Rate

Our analysis has revealed the factors that affect the value of the equilibrium exchange rate. Now we use this analysis to take a closer look at the response of the exchange rate to changes in interest rates and money growth.

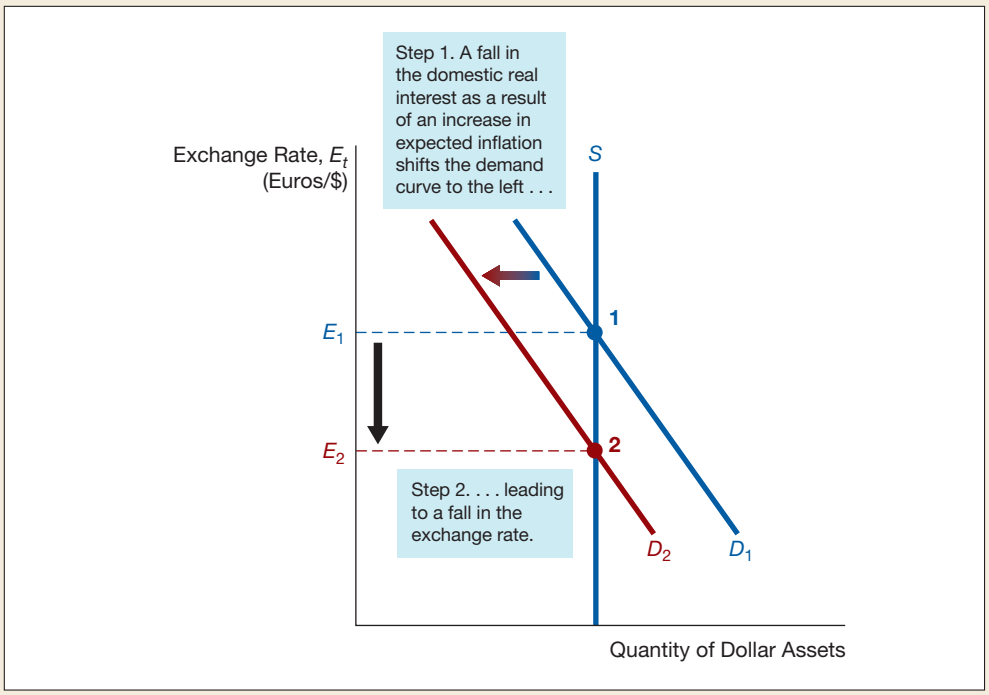
Changes in domestic interest rates i^D are often cited as a major factor affecting exchange rates. For example, we see headlines in the financial press like this one: "Dollar Recovers as Interest Rates Edge Upward." But is the positive correlation suggested in this headline true in every case?

Mini-lecture

FIGURE 18-6

Effect of a Rise in the Domestic Interest Rate as a Result of an Increase in Expected Inflation

Because a rise in domestic expected inflation leads to a decline in expected dollar appreciation that is larger than the increase in the domestic interest rate, the relative expected return on domestic (dollar) assets falls. The demand curve shifts to the left, and the equilibrium exchange rate falls from E_1 to E_2.


Step 1. A fall in the domestic real interest as a result of an increase in expected inflation shifts the demand curve to the left . . .
Exchange Rate, E_t (Euros/$)
S
E_1
1
E_2
2
Step 2. . . . leading to a fall in the exchange rate.
D_2
D_1
Quantity of Dollar Assets

Not necessarily, because to analyze the effects of interest rate changes, we must carefully distinguish the sources of the changes. The Fisher equation (Chapter 4) states that a nominal interest rate such as i^D equals the *real* interest rate plus expected inflation: $i = r + \pi^e$. The Fisher equation thus indicates that the interest rate i^D can change for two reasons: Either the real interest rate r changes or the expected inflation rate π^e changes. The effect on the exchange rate is quite different, depending on which of these two factors is the source of the change in the nominal interest rate.

Suppose the domestic real interest rate increases so that the nominal interest rate i^D rises while expected inflation remains unchanged. In this case, it is reasonable to assume that the expected appreciation of the dollar will be unchanged because expected inflation is unchanged. In this case, the increase in i^D increases the relative expected return on dollar assets, raises the quantity of dollar assets demanded at each level of the exchange rate, and shifts the demand curve to the right. We end up with the situation depicted in Figure 18-3, which analyzes an increase in i^D, holding everything else constant. Our model of the foreign exchange market produces the following result: *When domestic real interest rates rise, the domestic currency appreciates.*

When the nominal interest rate rises because of an increase in expected inflation, we get a different result from the one shown in Figure 18-3. The rise in expected domestic inflation leads to a decline in the expected appreciation of the dollar, which is typically found to be larger than the increase in the domestic interest rate i^D. As a result, at any given exchange rate, the relative expected return on domestic (dollar) assets falls, the demand curve shifts to the left, and the exchange rate falls from E_1 to E_2, as shown in Figure 18-6. Our analysis leads to this conclusion: ***When domestic interest rates rise due to an expected increase in inflation, the domestic currency depreciates.***

Because this result is completely different from the result that is obtained when the rise in the domestic interest rate is associated with a higher real interest rate, we must always distinguish between *real* and *nominal* measures when analyzing the effects of interest rates on exchange rates.

APPLICATION The Global Financial Crisis and the U.S. Dollar

With the start of the global financial crisis in August 2007, the U.S. dollar began an accelerated decline in value, falling by 9% against the euro until mid-July of 2008. After hitting an all-time low against the euro on July 11, the value of the U.S. dollar suddenly shot upward, by over 20% against the euro by the end of October. What is the relationship between the global financial crisis and these large swings in the value of the U.S. dollar?

Supply and demand analysis of what happened to the dollar is illustrated in Figure 18-7. Because we are analyzing what happened to the value of the dollar, we

Mini-lecture

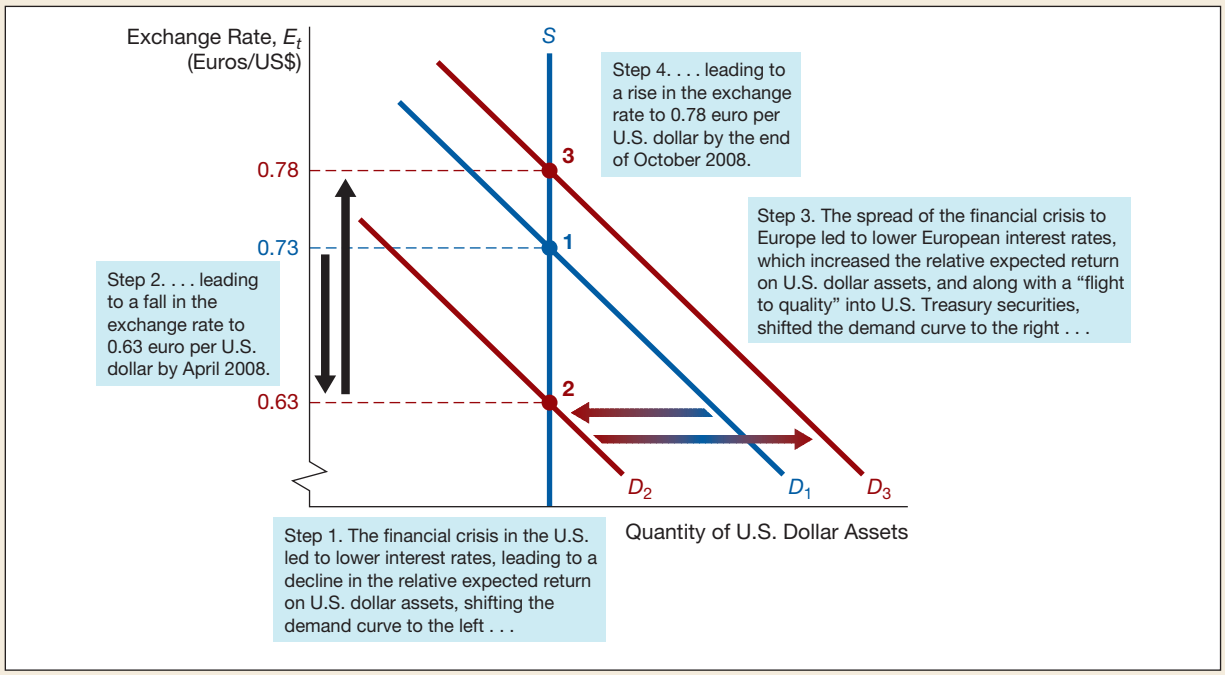

Exchange Rate, E_t (Euros/US$)
S
0.78
0.73
0.63
3
1
2
D_2
D_1
D_3
Quantity of U.S. Dollar Assets
Step 1. The financial crisis in the U.S. led to lower interest rates, leading to a decline in the relative expected return on U.S. dollar assets, shifting the demand curve to the left . . .
Step 2. . . . leading to a fall in the exchange rate to 0.63 euro per U.S. dollar by April 2008.
Step 3. The spread of the financial crisis to Europe led to lower European interest rates, which increased the relative expected return on U.S. dollar assets, and along with a "flight to quality" into U.S. Treasury securities, shifted the demand curve to the right . . .
Step 4. . . . leading to a rise in the exchange rate to 0.78 euro per U.S. dollar by the end of October 2008.

FIGURE 18-7 Global Financial Crisis and the U.S. Dollar

The financial crisis in the United States led to lower interest rates, leading to a decline in the relative expected return on U.S. dollar assets, which shifted the demand curve to the left and lowered the exchange rate to 0.63 euro per U.S. dollar by April 2008. The spread of the global financial crisis to Europe led to lower European interest rates, which increased the relative expected return of U.S. dollar assets, and along with a "flight to quality" to U.S. Treasury securities, shifted the demand curve to the right and raised the exchange rate to 0.78 euro per U.S. dollar by the end of October 2008.

are treating the U.S. dollar as the domestic currency. The exchange rate on the vertical axis in the diagram is therefore quoted as euros per U.S. dollar, so a rise in the exchange rate is an appreciation of the U.S. dollar. The horizontal axis is the quantity of U.S. (dollar) assets. In August 2007, before the crisis started, the exchange rate for the U.S. dollar was 0.73 euro per dollar at equilibrium point 1.

During 2007, the negative effects of the financial crisis on economic activity were mostly confined to the United States. The Federal Reserve acted aggressively to lower interest rates to counter the contractionary effects, decreasing the federal funds rate target by 325 basis points from September 2007 to April 2008. In contrast, other central banks like the ECB did not see the need to lower interest rates, particularly because high energy prices had led to a surge in inflation. The relative expected return on (U.S.) dollar assets thus declined, shifting the demand curve for dollar assets to the left in Figure 18-7, from D_1 to D_2. The equilibrium moved from point 1 in August 2007 to point 2 in April 2008, leading to a decline in the equilibrium exchange rate to 0.63 euro per U.S. dollar. Our analysis of the foreign exchange market thus explains why the early phase of the global financial crisis led to a decline in the value of the U.S. dollar.

We now turn to the rise in the value of the U.S. dollar. Starting in the summer of 2008, the effects of the global financial crisis on economic activity began to spread more widely throughout the world. The European Central Bank started to cut interest rates, with the expectation that further rate cuts would follow, as indeed did occur. The expected decline in foreign interest rates then increased the relative expected return on U.S. dollar assets, leading to a rightward shift in the demand curve from D_2 to D_3, with the equilibrium moving to point 3 where the U.S. dollar rose to 0.78 euro per U.S. dollar by the end of October 2008. Another factor driving the U.S. dollar upward was the "flight to quality" that occurred when the global financial crisis reached a particularly virulent stage in September and October. Both Americans and foreigners now wanted to put their money into the safest assets possible: U.S. Treasury securities. The resulting increase in the demand for U.S. dollar assets provided an additional reason for the demand curve for U.S. dollar assets to shift out to the right, thereby helping to produce a sharp appreciation of the U.S. dollar. ◆

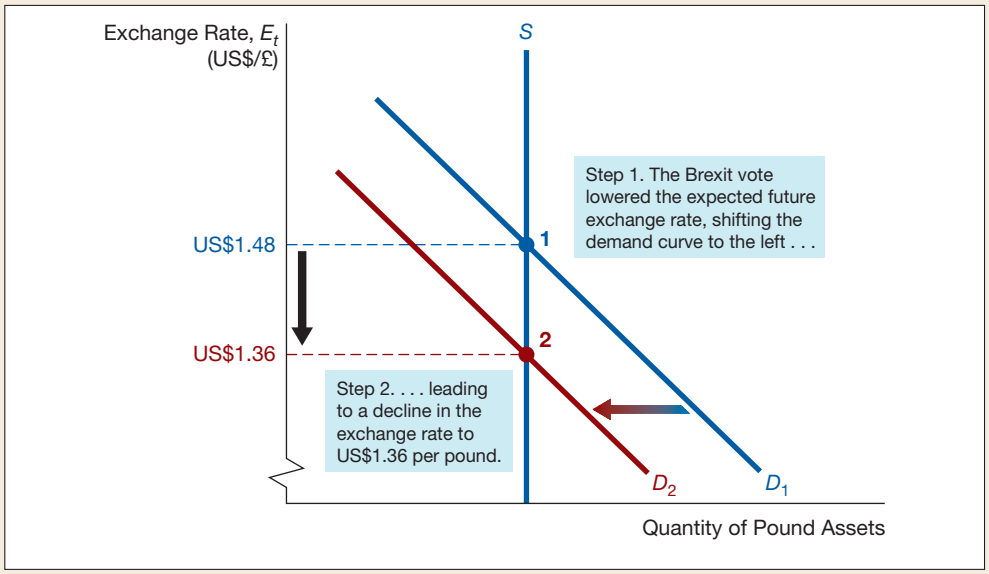
APPLICATION Brexit and the British Pound

As noted in the introduction, the Brexit vote in the United Kingdom on June 23, 2016, led to nearly a 10% depreciation in the British pound, from US$1.48 to the pound on June 23, just before the vote, to US$1.36 per pound on June 24. What explains the large one-day decline in the exchange rate for the pound?

Here we use our supply and demand analysis of the foreign exchange market in Figure 18-8 to answer this question. Because we are analyzing what happened to the value of the British pound, we are treating the pound as the domestic currency and thus are looking at the supply of and demand for pounds. The exchange rate on the vertical axis in the figure is therefore quoted as U.S. dollars per pound, so a rise in the exchange rate is an appreciation of the pound. The horizontal axis is the quantity of British pound assets. On June 23, right before the Brexit vote, the initial equilibrium was at point 1, with the equilibrium exchange rate for the pound at US$1.48 per pound.

Mini-lecture

FIGURE 18-8
Brexit and the British Pound
The Brexit vote lowered the expected return on British pound assets, which shifted the demand curve to the left and lowered the exchange rate to US$1.36 per pound.


Exchange Rate, E_t (US$/£)
S
Step 1. The Brexit vote lowered the expected future exchange rate, shifting the demand curve to the left . . .
US$1.48
1
US$1.36
2
Step 2. . . . leading to a decline in the exchange rate to US$1.36 per pound.
D_2
D_1
Quantity of Pound Assets

The Brexit vote would result in the United Kingdom no longer having access to the "one market" of the European Union, thereby likely erecting significant trade barriers to the export of British goods and services, especially financial services, which is one of Britain's most important industries. With higher expected European Union trade barriers, the demand for British goods and services would fall in the future, so the expected value of the pound would be lower in the future. The relative expected return on British (pound) assets therefore fell, and so the quantity demanded of pound assets declined at any given exchange rate, shifting the demand curve for pound assets to the left. The result was the sharp fall in the equilibrium exchange rate for the British pound to US$1.36 per pound. ◆

SUMMARY

1. Foreign exchange rates (the price of one country's currency in terms of another's) are important because they affect the price of domestically produced goods sold abroad and the cost of foreign goods bought domestically.
2. The theory of purchasing power parity suggests that long-run changes in the exchange rate between the currencies of two countries are determined by changes in the relative price levels in the two countries. Other factors that affect exchange rates in the long run are tariffs and quotas, import demand, export demand, and productivity.
3. In the short run, exchange rates are determined by changes in the relative expected return on domestic assets, which cause the demand curve to shift. Any factor that changes the relative expected return on domestic assets will lead to changes in the exchange rate. Such factors include changes in the interest rates on domestic and foreign assets as well as changes in any of the factors that affect the long-run exchange rate and hence the expected future exchange rate.
4. The asset market approach to exchange rate determination can explain the changes in the value of the dollar during the global financial crisis and the sharp decline in the value of the British pound after the Brexit vote.

MyLab Economics Make the grade with MyLab Economics: Select end-of-chapter questions can be found on MyLab Economics. You can practise them as often as you want, and most feature step-by-step guided instructions to help you find the right answer.

QUESTIONS

1. Suppose that you are considering going on vacation abroad and that the euro has appreciated by 15% with respect to the Canadian dollar. Would you be more or less willing to visit Rome and Paris?
2. "A country is always worse off when its currency is weak (falls in value)." Is this statement true, false, or uncertain? Explain your answer.
3. When the Canadian dollar depreciates, what happens to exports and imports in Canada?
4. If the Japanese price level rises by 5% relative to the price level in Canada, what does the theory of purchasing power parity predict will happen to the value of the Japanese yen in terms of dollars?
5. If the demand for a country's exports falls at the same time that tariffs on imports are raised, will the country's currency tend to appreciate or depreciate in the long run?
6. When the Bank of Canada conducts an expansionary monetary policy, what happens to the money supply? How does this affect the supply of dollar assets?
7. From 2009 to 2011, the economies of Australia and Switzerland suffered relatively mild effects from the global financial crisis. At the same time, many countries in the euro area were hit hard by high unemployment and burdened with unsustainably high government debts. How should this have affected the euro/Swiss franc and euro/Australian dollar exchange rates?
8. In the mid-to late 1970s, the yen appreciated in value relative to the U.S. dollar, even though Japan's inflation rate was higher than America's. How can this be explained by improvements in the productivity of Japanese industry relative to U.S. industry?
9. Suppose the governor of the Bank of Canada announces a new set of reforms that includes a new anti-inflation program. Assuming the announcement is believed by the public, what will happen to the exchange rate on the Canadian dollar?
10. If the Indian government unexpectedly announces that it will be imposing higher tariffs on foreign goods one year from now, what will happen to the value of the Indian rupee today?
11. If nominal interest rates in Canada rise but real interest rates fall, predict what will happen to the Canadian dollar exchange rate.
12. If American auto companies make a breakthrough in automobile technology and are able to produce a car that gets 200 miles to the gallon, what will happen to the U.S. dollar exchange rate?
13. If Mexicans go on a spending spree and buy twice as much French perfume and twice as many Japanese TVs, English sweaters, Swiss watches, and bottles of Italian wine, what will happen to the value of the Mexican peso?
14. Through the summer and fall of 2008, as the global financial crisis began to take hold, international financial institutions and sovereign wealth funds significantly increased their purchases of U.S. Treasury securities as a safe haven investment. How should this have affected U.S. dollar exchange rates?
15. In September 2012, the Federal Reserve announced a large-scale asset-purchase program (known as QE3) designed to lower intermediate and longer-term interest rates. What effect should this have had on the dollar/euro exchange rate?
16. On June 23, 2016, voters in the United Kingdom voted to leave the European Union. From June 16 to June 23, 2016, the exchange rate between the British pound and the U.S. dollar increased from 1.41 U.S. dollars per pound to 1.48 U.S. dollars per pound. What can you say about market expectations regarding the result of the referendum?

APPLIED PROBLEMS

17. A German sports car is selling for 70 000 euros. What is the dollar price in the United States of the German car if the exchange rate is 0.90 euro per dollar?

18. If the Canadian dollar to U.S. dollar exchange rate is 1.28 and the British pound to U.S. dollar exchange rate is 0.62, what must the Canadian dollar to British pound exchange rate be?

19. The New Zealand dollar to U.S. dollar exchange rate is 1.36, and the British pound to U.S. dollar exchange rate is 0.62. If you find that the British pound to New Zealand dollar is trading at 0.49, what can you do to earn a riskless profit?

20. In 1999, the euro was trading at US$0.90 per euro. If the euro is now trading at US$1.16 per euro, what is the percentage change in the euro's value? Is this an appreciation or a depreciation?

21. The Mexican peso is trading at 10 pesos per U.S. dollar. If the expected U.S. inflation rate is 2% while the expected Mexican inflation rate is 23% over the next year, given PPP, what is the expected exchange rate in one year?

22. If the price level recently increased by 20% in England while falling by 5% in Canada, by how much must the exchange rate change if PPP holds? Assume that the current exchange rate is 0.55 pound per dollar.

For Problems 23–25, use a graph of the foreign exchange market for dollars to illustrate the effects described in each problem.

23. If expected inflation drops in Europe so that interest rates fall there, what will happen to the exchange rate on the Canadian dollar?

24. If the European Central Bank decides to pursue a contractionary monetary policy to fight inflation, what will happen to the value of the Canadian dollar?

25. If a strike takes place in France, making it harder to buy French goods, what will happen to the value of the Canadian dollar?

DATA ANALYSIS PROBLEMS

1. **Real-Time Data Analysis** Go to the St. Louis Federal Reserve FRED database and find data on the exchange rate of U.S. dollars per British pound (DEXUSUK). A Mini Cooper can be purchased in London, England, for £17 865 or in Boston, United States, for $23 495.

a. Use the most recent exchange rate available to calculate the real exchange rate of the London Mini per Boston Mini.
b. Based on your answer to part (a), are Mini Coopers relatively more expensive in Boston or in London?
c. What price in British pounds would make the Mini Cooper equally expensive in both locations, all else being equal?

2. **Real-Time Data Analysis** Go to the St. Louis Federal Reserve FRED database and find data on the daily U.S. dollar exchange rates for the euro (DEXUSEU), British pound (DEXUSUK), and Japanese yen (DEXJPUS). Also find data on the daily three-month London Interbank Offer Rate, or LIBOR, for the United States dollar (USD-3MTD156N), euro (EUR3MTD156N), British pound (GBP3MTD156N), and Japanese yen (JPY3MTD156N). LIBOR is a measure of interest rates denominated in each country's respective currency.

a. Calculate the difference between the LIBOR rate in the United States and the LIBOR rates in the three other countries using the data from one year ago and the most recent data available.
b. Based on the changes in interest rate differentials, do you expect the dollar to depreciate or appreciate against the other currencies?
c. Report the percentage change in the exchange rates over the past year. Are the results you predicted in part (b) consistent with the actual exchange rate behaviour?

19 The International Financial System

Learning Objectives

19.1 Use graphs and T-accounts to illustrate the distinctions between the effects of sterilized and unsterilized interventions on foreign exchange markets.

19.2 Interpret the relationships among the current account and the financial account in the balance of payments.

19.3 Identify the mechanisms for maintaining a fixed exchange rate, and assess the challenges faced by fixed exchange rate regimes.

19.4 Summarize the advantages and disadvantages of capital controls.

19.5 Assess the role of the IMF as an international lender of last resort.

19.6 Identify the ways in which international monetary policy and exchange-rate arrangements can affect domestic monetary policy operations.

19.7 Summarize the advantages and disadvantages of exchange-rate targeting.

Preview

As the Canadian economy and the economies of the rest of the world grow more interdependent, a country's monetary policy can no longer be conducted without taking international considerations into account. In this chapter, we examine how international financial transactions and the structure of the international financial system affect monetary policy. We also examine the evolution of the international financial system during the past half century and consider where it may be headed in the future.

19.1 INTERVENTION IN THE FOREIGN EXCHANGE MARKET

LO 19.1 Use graphs and T-accounts to illustrate the distinctions between the effects of sterilized and unsterilized interventions on foreign exchange markets.

In Chapter 18, we analyzed the foreign exchange market as if it were a completely free market that responds to all conventional market pressures. Like many other markets, however, the foreign exchange market is not free of government intervention; central banks regularly engage in international financial transactions, the purchase and sale of their currencies, called **foreign exchange interventions**, to influence exchange rates. In our current international environment, exchange rates fluctuate from day to day, but central banks attempt to influence their countries' exchange rates by buying and selling currencies. We can use the exchange-rate analysis we developed in Chapter 18 to explain the impact of central bank intervention on the foreign exchange market.

Foreign Exchange Intervention and the Money Supply

The first step in understanding how central bank intervention in the foreign exchange market affects exchange rates is to explore the impact on the monetary base of a central bank sale in the foreign exchange market of some of its holdings of assets denominated in a foreign currency (called **international reserves**). Suppose the Bank of Canada decides to sell $1 billion of its foreign assets in exchange for $1 billion of Canadian currency. The Bank of Canada's purchase of dollars has two effects. First, it reduces the Bank's holdings of international reserves by $1 billion. Second, because the Bank's purchase of currency removes it from the hands of the public, currency

in circulation falls by $1 billion. We can see this in the following T-account for the Bank of Canada:

Bank of Canada

Assets		Liabilities	
Foreign assets (international reserves)	−$1 billion	Currency in circulation	−$1 billion

Because the monetary base is made up of currency in circulation plus reserves, this decline in currency implies that the monetary base has fallen by $1 billion.

If, as is more likely, the persons buying the foreign assets pay for them with cheques written on accounts at domestic banks rather than with currency, then the Bank of Canada deducts the $1 billion from the reserve deposits it holds for these banks. The result is that deposits with the Bank of Canada (reserves) decline by $1 billion, as shown in the following T-account:

Bank of Canada

Assets		Liabilities	
Foreign assets (international reserves)	−$1 billion	Deposits with the Bank of Canada (reserves)	−$1 billion

In this case, the outcome of the Bank of Canada sale of foreign assets and the purchase of dollar deposits is a $1 billion decline in reserves and, as before, a $1 billion decline in the monetary base, because reserves are also a component of the monetary base.

We now see that the outcome for the monetary base is exactly the same when a central bank sells foreign assets to purchase domestic bank deposits or domestic currency. This is why, when we say that a central bank has purchased its domestic currency, we do not have to distinguish between currency and bank deposits denominated in the domestic currency. We have thus reached an important conclusion: ***A central bank's purchase of domestic currency and corresponding sale of foreign assets in the foreign exchange market leads to an equal decline in its international reserves and the monetary base.***

We could have reached the same conclusion by a more direct route. A central bank sale of a foreign asset is no different from an open market sale of a government bond. We learned in our exploration of the money supply process that an open market sale leads to an equal decline in the monetary base; therefore, a sale of foreign assets also leads to an equal decline in the monetary base. By similar reasoning, a central bank purchase of foreign assets paid for by selling domestic currency, like an open market purchase, leads to an equal rise in the monetary base. Thus we reach the following conclusion: ***A central bank's sale of domestic currency to purchase foreign assets in the foreign exchange market results in an equal rise in its international reserves and the monetary base***.

The intervention we have just described, in which a central bank allows the purchase or sale of domestic currency to have an effect on the monetary base, is called an **unsterilized foreign exchange intervention**. But what if the central bank does not want the purchase or sale of domestic currency to affect the monetary base? All it has to do is counteract the effect of the foreign exchange intervention by conducting an

offsetting open market operation in the government bond market. For example, in the case of a $1 billion purchase of dollars by the Bank of Canada and a corresponding $1 billion sale of foreign assets, which, as we have seen, would decrease the monetary base by $1 billion, the Bank of Canada can conduct an open market purchase of $1 billion of government bonds, which will increase the monetary base by $1 billion. The resulting T-account for the foreign exchange intervention and the offsetting open market operation leaves the monetary base unchanged:

Bank of Canada

Assets		Liabilities	
Foreign assets (international reserves)	−$1 billion	Monetary base	0
Government bonds	+$1 billion		

A foreign exchange intervention with an offsetting open market operation that leaves the monetary base unchanged is called a **sterilized foreign exchange intervention**.

Now that we understand that there are two types of foreign exchange interventions—unsterilized and sterilized—let's look at how each affects the exchange rate.

Unsterilized Intervention

Intuition might lead you to suspect that if a central bank wants to raise the value of the domestic currency, it should buy its currency in the foreign exchange market and sell foreign assets. Indeed, this intuition is correct in the case of an unsterilized intervention.

Recall that in an unsterilized intervention, if the Bank of Canada decides to buy dollars and therefore sells foreign assets in exchange for dollar assets, this exchange works just like an open market sale of bonds to decrease the monetary base. Hence the purchase of dollars leads to a decrease in the money supply, which raises the domestic interest rate and increases the relative expected return on dollar assets. As a result, the demand curve shifts to the right from D_1 to D_2, as shown in Figure 19-1, and the exchange rate rises to E_2.[1]

Our analysis leads us to the following conclusion about unsterilized interventions in the foreign exchange market: ***An unsterilized intervention in which domestic currency is bought and foreign assets are sold leads to a fall in international reserves, a fall in the money supply, and an appreciation of the domestic currency***.

The reverse result is found for an unsterilized intervention in which domestic currency is sold and foreign assets are purchased. The sale of domestic currency and purchase of foreign assets (which increases international reserves) work like an open market purchase to increase the monetary base and the money supply. The increase in the money supply lowers the interest rate on dollar assets. The resulting decrease in the relative expected return on dollar assets means that people will buy fewer dollar assets, so the demand curve shifts to the left and the exchange rate falls. ***An unsterilized intervention in which domestic***

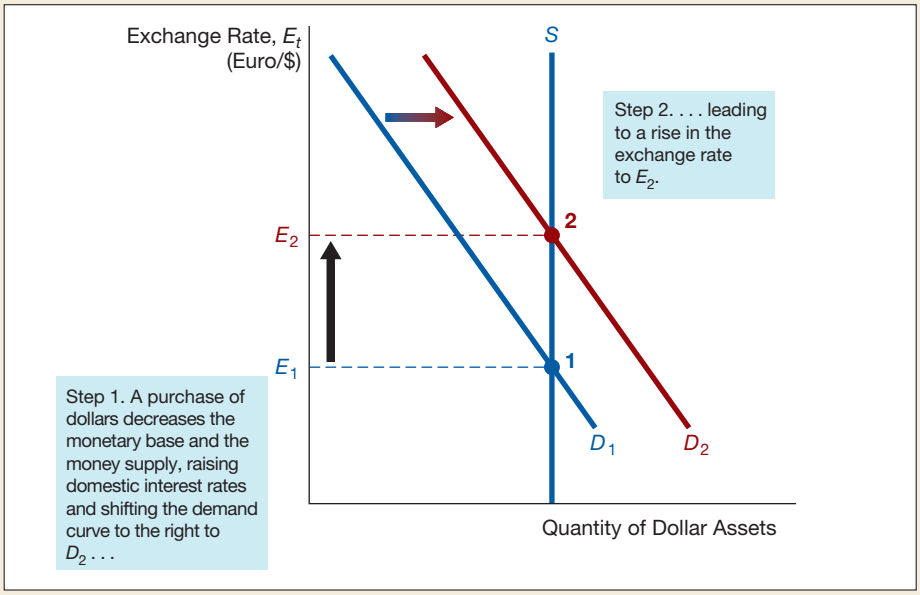
[1]An unsterilized intervention in which the Bank of Canada buys dollars decreases the amount of dollar assets slightly because it leads to a decrease in the monetary base while leaving the amount of government bonds in the hands of the public unchanged. The curve depicting the supply of dollar assets thus shifts to the left slightly, which also works toward raising the exchange rate, yielding the same conclusion derived in Figure 19-1. Because the resulting increase in the monetary base is only a minuscule fraction of the total amount of dollar assets outstanding, the supply curve shifts by an imperceptible amount. This is why Figure 19-1 is drawn with the supply curve unchanged.

Mini-lecture

FIGURE 19-1

Effect of an Unsterilized Purchase of Dollars and Sale of Foreign Assets

A purchase of dollars and the consequent open market sale of foreign assets decrease the monetary base and the money supply. The resulting fall in the money supply leads to a rise in domestic interest rates, which raises the relative expected return on dollar assets. The demand curve shifts to the right from D_1 to D_2, and the equilibrium exchange rate rises from E_1 to E_2.


Exchange Rate, E_t (Euro/$)
S
Step 2. . . . leading to a rise in the exchange rate to E_2.
E_2
2
E_1
1
Step 1. A purchase of dollars decreases the monetary base and the money supply, raising domestic interest rates and shifting the demand curve to the right to D_2 . . .
D_1
D_2
Quantity of Dollar Assets

currency is sold and foreign assets are purchased leads to a rise in international reserves, a rise in the money supply, and a depreciation of the domestic currency.

Sterilized Intervention

The key point to remember about a sterilized intervention is that the central bank engages in offsetting open market operations, and so there is no impact on the monetary base and the money supply. In the context of the model of exchange-rate determination we have developed here, it is a straightforward task to show that a sterilized intervention has almost *no effect* on the exchange rate. A sterilized intervention leaves the money supply unchanged and thus has no direct way of affecting interest rates.[2] Because the relative expected return on dollar assets is unaffected, the demand curve remains at D_1 in Figure 19-1, and the exchange rate remains unchanged at E_1.

At first it might seem puzzling that a central bank purchase or sale of domestic currency that is sterilized does not lead to a change in the exchange rate. A central bank–sterilized purchase of domestic currency cannot raise the exchange rate because, with no effect on the domestic money supply or interest rates, any resulting rise in the exchange rate would lead to an excess supply of dollar assets. With more people willing to sell dollar assets than willing to buy them, the exchange rate would fall back to its initial equilibrium level, where the demand and supply curves intersect.

[2]A sterilized intervention changes the amount of foreign securities relative to domestic securities in the hands of the public, called a *portfolio balance effect*. Through this effect, the central bank might be able to influence the interest differential between domestic and foreign assets, which in turn affects the relative expected return on domestic assets. Empirical evidence has not revealed this portfolio balance effect to be significant. However, a sterilized intervention *could* indicate what central banks want to happen to the future exchange rate and so might provide a signal about the course of future monetary policy and change the expected future exchange rate E^e_{t+1}. In this way, a sterilized intervention could lead to a shift in the demand curve for domestic assets and ultimately affect the exchange rate. However, the future change in monetary policy—not the sterilized intervention—is the source of the exchange-rate effect. For a further discussion of the signalling and portfolio balance effects and the possible differential effects of sterilized versus unsterilized intervention, see Paul Krugman, Maurice Obstfeld, and Mark Melitz, *International Economics: Theory and Policy*, 11th ed. (New York: Pearson, 2018).

19.2 BALANCE OF PAYMENTS

LO 19.2 Interpret the relationships among the current account and the financial account in the balance of payments.

Because international financial transactions strongly affect a country's monetary policy, it is worth knowing how these transactions are measured. The **balance of payments** is a bookkeeping system that records the transactions between a nation (private sector and government) and foreign countries. Here we examine the key items in the balance of payments, which you often hear about in the media.

Current Account

The **current account** shows a country's current international transactions (that is, transactions that do not involve the purchase or sale of financial assets) for a given year. The current account is composed of three items: the trade balance, net investment income, and transfers.

The **trade balance on goods and services** (often called simply the **trade balance**) equals the total value of exports of goods and services minus the total value of imports of goods and services. (The trade balance is also referred to as **net exports**.) When merchandise imports are greater than exports (as they were by $36.8 billion in 2019 in Canada), the country runs a trade deficit. If exports of goods and services exceed imports, the country runs a trade surplus.

In 2019, *net investment income* was −$2.2 billion for Canada because Canadians paid out more investment income than they received from abroad. *Transfers* are funds sent by domestic residents and the government to foreigners; they include *remittances* (money sent by domestic workers to foreigners, usually relatives), pensions, and foreign aid. Because Canadians made more transfers to foreigners than foreigners made to Canada, transfers were −$8.4 billion in 2019. The sum of these two items plus the trade balance (net exports) is the **current account balance**, which in 2019 had a deficit of −$41.4 billion:

$$-\$47.4 \text{ billion} = (-\$36.8 \text{ billion}) + (-\$2.2 \text{ billion}) + (-\$8.4 \text{ billion})$$

$$\text{Current account balance} = \text{trade balance} + \text{net investment income} + \text{transfers}$$

When a country is running a current account deficit, as Canada did, then it is making more payments to foreigners on a current basis than it is receiving from foreigners. Like any deficit, a current account deficit has to be financed. To see how, we now turn to the financial account.

Financial Account

The **financial account** shows the international transactions that involve the purchase or sale of assets. When Canadians increase their net holdings of foreign assets, this change is recorded as a *net Canadian acquisition of financial assets* in the financial account. When foreigners increase their net holdings of Canadian assets, this change is recorded as a *net Canadian incurrence of liabilities* because foreigners' claims on Canadian assets are liabilities for Canadians. The difference between the *net Canadian acquisition of financial assets* and *net Canadian incurrence of liabilities* is the **financial account balance**. In 2019, the financial account balance plus other items, which for practical purposes we will include with the financial account balance, was −$51.2 billion. This financial account balance tells us that in 2019, Canada experienced an increase of liabilities to foreigners that exceeded the increase in Canadian foreign asset holdings by $51.2 billion. In other words, our Canadian indebtedness to foreigners increased by $51.2 billion.

Global **Should We Worry About the Large U.S. Current Account Deficit?**

The large U.S. current account deficit in recent years—in 2019 it was US$480 billion, 2.2% of GDP—raises concerns for several reasons. First, because a large part of the deficit comes from a trade deficit, we can conclude that at current exchange rates, foreigners' demand for U.S. exports is far less than Americans' demand for imports. Our supply and demand analysis of the foreign exchange market suggests that low demand for U.S. exports and high U.S. demand for imports may lead to a decline in the value of the U.S. dollar.

Second, a large current account deficit means that the United States is becoming more indebted to foreigners. This borrowing will have to be paid back at some point in the future. When the bill comes due, Americans will not be as rich as they otherwise would be.

However, a large current account deficit indicates large capital inflows into the United States. These capital inflows might be funding highly attractive investment opportunities in the United States. These investments might increase U.S. wealth so much in the future that even when foreigners are paid back, Americans, on net, are richer.

The financial account balance tells us the amount of *net capital inflows* into Canada. When foreigners increase their holdings of Canadian assets, capital has flowed into Canada. In contrast, when Canadians increase their holdings of foreign assets, capital has flowed out of Canada. So, in 2019, foreigners increased their holdings of Canadian assets by $51.2 billion more than Canadians increased their holdings of foreign assets. The result was a net capital inflow of $51.2 billion into Canada.

You probably have noticed that this financial account balance is equal in absolute value to the current account deficit. This equality makes sense because if Canadians are spending more than they are taking in on the current account, they must be financing this spending by borrowing an equal amount from foreigners. Another way to see this equality is to use the accounting principle that the uses of funds must equal the sources of funds. Hence the current account deficit, which provides the net uses of funds for current items, must equal the sources of funds, which is what the financial account balance tells us. In the United States, the large current account deficit in recent years, which in 2019 was over US$480 billion, has raised serious concerns that these large deficits may have negative consequences for the U.S. economy (see the Global box "Should We Worry About the Large U.S. Current Account Deficit?").

19.3 EXCHANGE RATE REGIMES IN THE INTERNATIONAL FINANCIAL SYSTEM

LO 19.3 Identify the mechanisms for maintaining a fixed exchange rate, and assess the challenges faced by fixed exchange rate regimes.

Exchange rate regimes in the international financial system are classified into two basic types: fixed and floating. In a **fixed exchange rate regime**, the value of a currency is pegged relative to the value of another currency (called the **anchor currency**) so that the exchange rate is fixed in terms of the anchor currency. In a **floating exchange rate regime**, the value of a currency is allowed to fluctuate against all other currencies. When countries intervene in foreign exchange markets in an attempt to influence their exchange rates by buying and selling foreign assets, the regime is referred to as a **managed float regime** (or a **dirty float**).

In examining past exchange rate regimes, we start with the gold standard of the late nineteenth and early twentieth centuries.

Gold Standard

Before World War I, the world economy operated under the **gold standard**, a fixed exchange rate regime in which the currencies of most countries were convertible directly into gold at fixed rates so that exchange rates between currencies were also fixed. One Canadian dollar bill, for example, could be exchanged for approximately 1/20 ounce of gold. Likewise, the British Treasury would exchange ¼ ounce of gold for £1 sterling. Because a Canadian could convert $20 into one ounce of gold, which could be used to buy £4, the exchange rate between the pound and the dollar was effectively fixed at $5 to the pound. The fixed exchange rates under the gold standard had the important advantage of encouraging world trade by eliminating the uncertainty that occurs when exchange rates fluctuate.

One problem with the gold standard was that monetary policy throughout the world was greatly influenced by the production of gold and gold discoveries. When gold production was low in the 1870s and 1880s, the money supply throughout the world grew slowly and did not keep pace with the growth of the world economy. The result was deflation (falling price levels). Gold discoveries in Alaska and South Africa in the 1890s greatly expanded gold production, causing money supplies to increase rapidly and price levels to rise (inflation) until World War I.

The Bretton Woods System

After World War II, the victors set up a fixed exchange rate system that became known as the **Bretton Woods system**, after the New Hampshire town in which the agreement was negotiated in 1944. The Bretton Woods system remained in effect until 1971.

The Bretton Woods agreement created the **International Monetary Fund (IMF)**, headquartered in Washington, DC, which had 30 original member countries in 1945 and currently has over 180. The IMF was given the task of promoting the growth of world trade by setting rules for the maintenance of fixed exchange rates and by making loans to countries that were experiencing balance-of-payments difficulties. In addition to monitoring the compliance of member countries with its rules, the IMF also took on the job of collecting and standardizing international economic data.

The Bretton Woods agreement also set up the International Bank for Reconstruction and Development, commonly referred to as the **World Bank**. Headquartered in Washington, DC, it provides long-term loans to help developing countries build dams, roads, and other physical capital that will contribute to their economic development. The funds for these loans are obtained primarily by the issuance of World Bank bonds, which are sold in the capital markets of the developed countries. In addition, the General Agreement on Tariffs and Trade (GATT), headquartered in Geneva, Switzerland, was set up to monitor rules for the conduct of trade between countries (tariffs and quotas). The GATT has since evolved into the **World Trade Organization (WTO)**.

Because the United States emerged from World War II as the world's largest economic power, with over half of the world's manufacturing capacity and the greater part of the world's gold, Bretton Woods put in place a system of fixed exchange rates that was based on the convertibility of U.S. dollars into gold (for foreign governments and central banks only) at $35 per ounce. The fixed exchange rates were to be maintained by intervention in the foreign exchange market by central banks in countries besides

the United States that bought and sold dollar assets, which these countries held as international reserves. The U.S. dollar, which was used by other countries to denominate the assets that they held as international reserves, was called the **reserve currency**. Thus an important feature of the Bretton Woods system was the establishment of the United States as the reserve currency country. Even after the breakup of the Bretton Woods system, the U.S. dollar kept its position as the reserve currency in which most international financial transactions are conducted.

The fixed exchange rate, which was a feature of the Bretton Woods system, was finally abandoned in 1973. From 1979 to 1990, the European Union instituted among its members its own fixed exchange rate system, called the European Monetary System (EMS). In theory, in the *exchange rate mechanism* (*ERM*) of this system, the exchange rate between any pair of currencies of the participating countries was not supposed to fluctuate outside narrow limits, called the "snake." In practice, all of the countries in the EMS pegged their currencies to the German mark.

How a Fixed Exchange Rate Regime Works

Figure 19-2 uses the supply and demand analysis of the foreign exchange market that we learned in the previous chapter to show how a fixed exchange rate regime works in practice. Panel (a) describes a situation in which the domestic currency is fixed relative to an anchor currency at E_{par}, while the demand curve has shifted left to D_1, perhaps

Mini-lecture

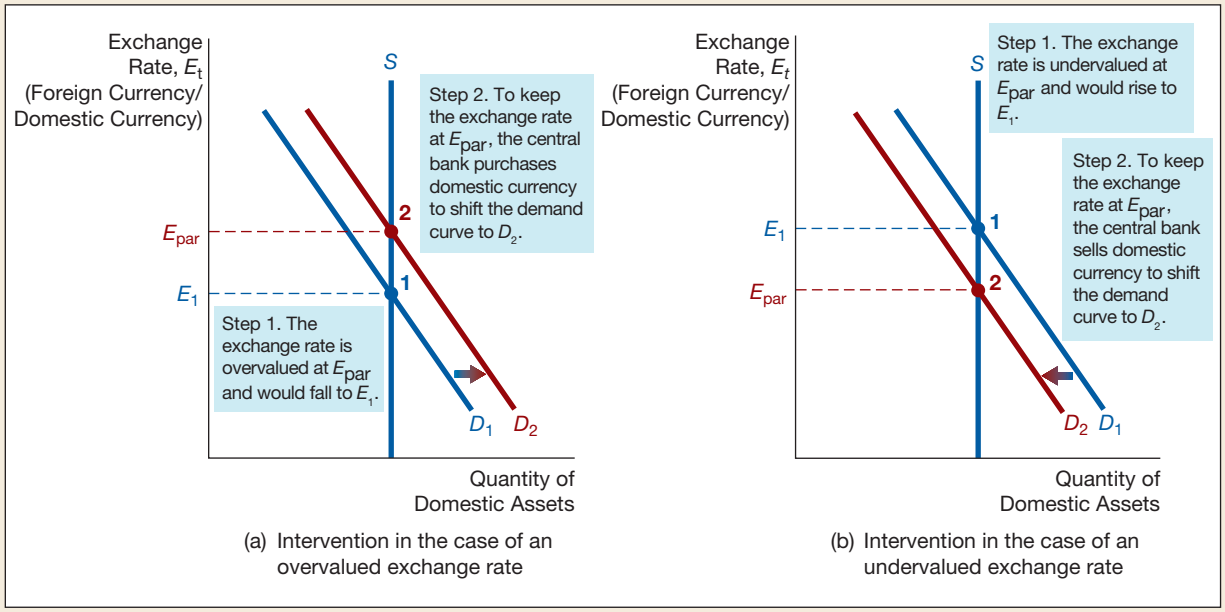

Exchange Rate, E_t (Foreign Currency/ Domestic Currency)
S
Step 2. To keep the exchange rate at E_{par}, the central bank purchases domestic currency to shift the demand curve to D_2.
E_{par}
E_1
2
1
Step 1. The exchange rate is overvalued at E_{par} and would fall to E_1.
D_1
D_2
Quantity of Domestic Assets
(a) Intervention in the case of an overvalued exchange rate
Exchange Rate, E_t (Foreign Currency/ Domestic Currency)
S
Step 1. The exchange rate is undervalued at E_{par} and would rise to E_1.
Step 2. To keep the exchange rate at E_{par}, the central bank sells domestic currency to shift the demand curve to D_2.
E_1
E_{par}
1
2
D_2
D_1
Quantity of Domestic Assets
(b) Intervention in the case of an undervalued exchange rate

FIGURE 19-2 Intervention in the Foreign Exchange Market Under a Fixed Exchange Rate Regime

In panel (a), the exchange rate at E_{par} is overvalued. To keep the exchange rate at E_{par} (point 2), the central bank must purchase domestic currency to shift the demand curve to D_2. In panel (b), the exchange rate at E_{par} is undervalued, so the central bank must sell domestic currency to shift the demand curve to D_2 and keep the exchange rate at E_{par} (point 2).

because foreign interest rates have risen, thereby lowering the relative expected return on domestic assets. At E_{par}, the exchange rate is now *overvalued*: The demand curve D_1 intersects the supply curve at an exchange rate E_1 that is lower than the fixed (par) value of the exchange rate E_{par}. To keep the exchange rate at E_{par}, the central bank must intervene in the foreign exchange market and purchase domestic currency by selling foreign assets. This action, like an open market sale, causes both the monetary base and the money supply to decrease, driving up the interest rate on domestic assets, i^D.[3] This increase in the domestic interest rate raises the relative expected return on domestic assets, shifting the demand curve to the right. The central bank will continue to purchase domestic currency until the demand curve reaches D_2 and the equilibrium exchange rate is at E_{par}, at point 2 in panel (a).

We thus conclude that ***when the domestic currency is overvalued, the central bank must purchase domestic currency to keep the exchange rate fixed, but as a result it loses international reserves***.

Panel (b) in Figure 19-2 describes the situation in which the demand curve has shifted to the right to D_1 because the relative expected return on domestic assets has risen and hence the exchange rate at E_{par} is *undervalued*: The initial demand curve D_1 intersects the supply curve at exchange rate E_1, which is above E_{par}. In this situation, the central bank must sell domestic currency and purchase foreign assets. This action works like an open market purchase; it increases the money supply and lowers the interest rate on domestic assets i^D. The central bank keeps selling domestic currency and lowering i^D until the demand curve shifts all the way to D_2, where the equilibrium exchange rate is at E_{par}—point 2 in panel (b). Our analysis thus leads us to the following result: ***When the domestic currency is undervalued, the central bank must sell domestic currency to keep the exchange rate fixed, but as a result, it gains international reserves***.

Devaluation and Revaluation As we have seen, if a country's currency is overvalued, its central bank's attempts to keep the currency from depreciating will result in a loss of international reserves. If the country's central bank eventually runs out of international reserves, it cannot keep its currency from depreciating, and a **devaluation**, in which the par exchange rate is reset at a lower level, must occur.

If, by contrast, a country's currency is undervalued, its central bank's intervention to keep the currency from appreciating leads to a gain of international reserves. As we will see shortly, the central bank might not want to acquire these international reserves, and so it might want to reset the par value of its exchange rate at a higher level (a **revaluation**).

Perfect Capital Mobility If perfect **capital mobility** exists—that is, if there are no barriers to domestic residents purchasing foreign assets or to foreigners purchasing domestic assets—then a sterilized exchange-rate intervention cannot keep the exchange rate at E_{par} because, as we saw earlier in the chapter, the relative expected return on domestic assets is unaffected. For example, if the exchange rate is overvalued, a sterilized purchase of domestic currency will leave the relative expected return and the demand curve unchanged, so pressure for a depreciation of the domestic currency is not removed. If the central bank keeps purchasing its domestic currency but continues to sterilize, it will just keep losing international reserves until it finally runs out of them and is forced to let the value of the currency seek a lower level.

[3]Because the exchange rate will continue to be fixed at E_{par}, the expected future exchange rate remains unchanged and so is not addressed in the analysis.

Speculative Attacks

A serious shortcoming of fixed exchange rate systems such as the Bretton Woods system and the European Monetary System is that such systems can lead to foreign exchange crises involving a **speculative attack** on a currency. A speculative attack involves massive sales of a weak currency or purchases of a strong currency that cause a sharp change in the exchange rate. In the following application, we use our model of exchange-rate determination to understand how the September 1992 exchange-rate crisis that rocked the European Monetary System came about.

APPLICATION The Foreign-Exchange Crisis of September 1992

In the aftermath of German reunification in October 1990, the German central bank, the Bundesbank, faced rising inflationary pressures, with inflation having accelerated from below 3% in 1990 to near 5% by 1992. To get monetary growth under control and to dampen inflation, the Bundesbank raised German interest rates to near double-digit levels. Figure 19-3 shows the consequences of these actions by the Bundesbank in the foreign

Mini-lecture

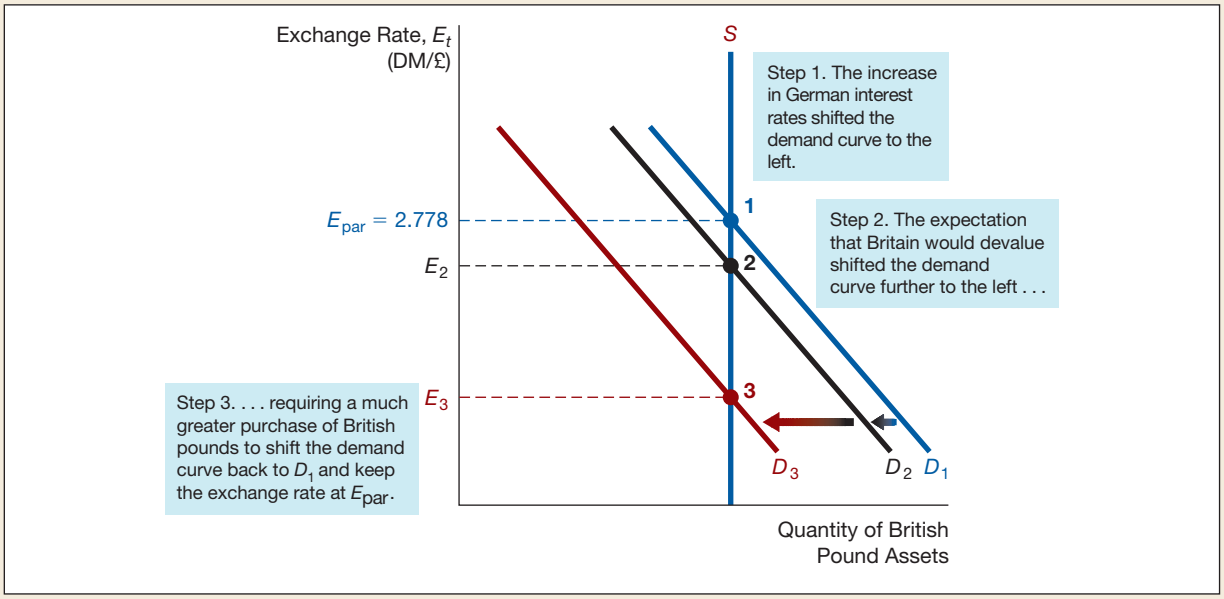

Exchange Rate, E_t (DM/£)
S
Step 1. The increase in German interest rates shifted the demand curve to the left.
E_{par} = 2.778
1
Step 2. The expectation that Britain would devalue shifted the demand curve further to the left . . .
E_2
2
Step 3. . . . requiring a much greater purchase of British pounds to shift the demand curve back to D_1 and keep the exchange rate at E_{par}.
E_3
3
D_3
D_2
D_1
Quantity of British Pound Assets

FIGURE 19-3 Foreign Exchange Market for British Pounds in 1992

The realization by speculators that the United Kingdom would soon devalue the pound decreased the relative expected return on British pound assets, resulting in a leftward shift of the demand curve from D_2 to D_3. As a result, the British central bank was forced to purchase a large amount of pounds in an effort to raise the interest rate so that the demand curve would shift back to D_1 and the exchange rate E_{par} would remain at 2.778 German marks per pound.

exchange market for British pounds. Note that in the diagram, the pound is the domestic currency and the German mark (deutsche mark, DM, Germany's currency before the advent of the euro in 1999) is the foreign currency.

The increase in German interest rates i^F lowered the relative expected return on British pound assets and shifted the demand curve to D_2, as shown in Figure 19-3. The intersection of the supply-and-demand curves at point 2 was now below the lower exchange-rate limit at that time (2.778 marks per pound, denoted E_{par}). To increase the value of the pound relative to the mark and to restore the mark/pound exchange rate to within the exchange rate mechanism limits, one of two things had to happen. In the first option, the Bank of England would have to pursue a contractionary monetary policy, thereby raising British interest rates sufficiently to shift the demand curve back to D_1 so that the equilibrium would remain at point 1 and the exchange rate would remain at E_{par}. Alternatively, the Bundesbank would have to pursue an expansionary monetary policy, thereby lowering German interest rates. Lower German interest rates would raise the relative expected return on British assets and shift the demand curve back to D_1 so that the exchange rate would remain at E_{par}.

The catch was that the Bundesbank, whose primary goal was to fight inflation, was unwilling to pursue an expansionary monetary policy, and the British, who were facing their worst recession in the postwar period, were unwilling to pursue a contractionary monetary policy to prop up the pound. This impasse became clear when, in response to great pressure from other members of the EMS, the Bundesbank was willing to lower its lending rates by only a token amount on September 14 after a speculative attack was mounted on the currencies of the Scandinavian countries. So, at some point in the near future, the value of the pound would have to decline to point 2. Speculators now knew that depreciation of the pound was imminent, and so the expected future exchange rate, E^e_{t+1}, fell. As a result, the relative expected return on the pound fell sharply, shifting the demand curve left to D_3 in Figure 19-3.

As a result of the large leftward shift of the demand curve, a huge excess supply of pound assets now existed at the par exchange rate E_{par}, which caused a massive sell-off of pounds (and purchases of marks) by speculators. The need for the British central bank to intervene to raise the value of the pound now became much greater and required a huge rise in British interest rates. A major intervention effort on the part of the Bank of England, which led to a rise in its lending rate from 10% to 15%, still wasn't enough to save the pound. The British were finally forced to give up on September 16: They pulled out of the ERM indefinitely and allowed the pound to depreciate by 10% against the mark.

Speculative attacks on other currencies forced devaluation of the Spanish peseta by 5% and of the Italian lira by 15%. To defend its currency, the Swedish central bank was forced to raise its daily lending rate to the astronomical level of 500%! By the time the crisis was over, the British, French, Italian, Spanish, and Swedish central banks had intervened to the tune of $100 billion; the Bundesbank alone had laid out $50 billion for foreign exchange intervention.

The attempt to prop up the European Monetary System was not cheap for these central banks. It is estimated that they lost $4 to $6 billion as a result of exchange-rate intervention during the crisis. What the central banks lost, the speculators gained. A speculative fund run by George Soros ran up $1 billion of profits during the crisis, and Citibank traders reportedly made $200 million. When an exchange-rate crisis arises, life certainly can be sweet for exchange-rate speculators. ◆

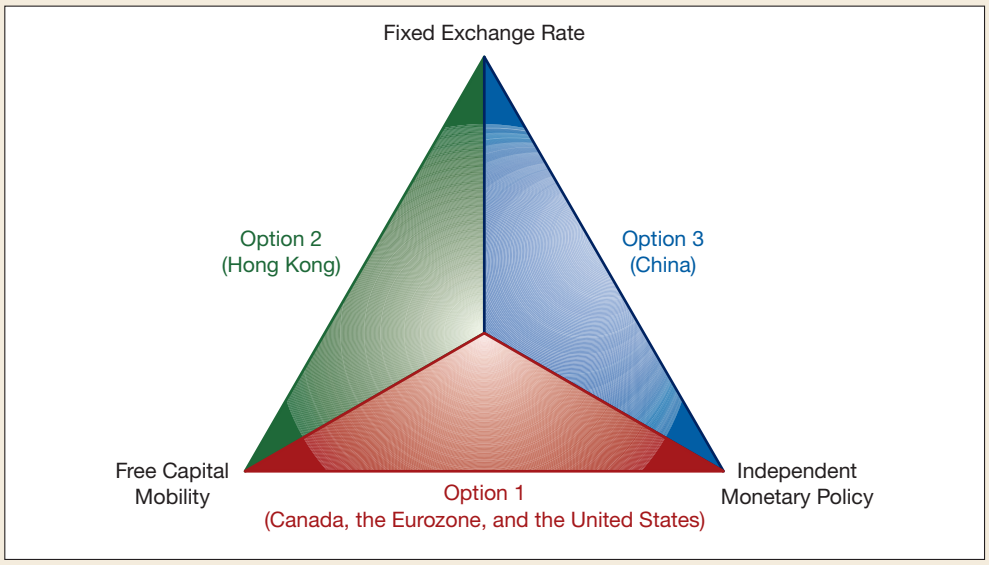
The Policy Trilemma

An important implication of the foregoing analysis is that a country that ties its exchange rate to an anchor currency of a larger country loses control of its monetary policy. If the larger country pursues a more contractionary monetary policy and raises interest rates, this action will lead to lower expected inflation in the larger country, thus causing an appreciation of the larger country's currency and a depreciation of the smaller country's currency. The smaller country, having locked its exchange rate to the anchor currency, will now find its currency overvalued and will therefore have to sell the anchor currency and buy its own to keep its currency from depreciating. The result of this foreign exchange intervention will be a decline in the smaller country's international reserves, a contraction of its monetary base, and a rise in interest rates. Sterilization of this foreign exchange intervention is not an option because this course of action would just lead to a continuing loss of international reserves until the smaller country is forced to devalue its currency. The smaller country no longer controls its monetary policy because movements in its interest rates are completely determined by movements in the larger country's interest rates.

Our analysis therefore indicates that a country (or a monetary union like the Eurozone) cannot pursue the following three policies at the same time: (1) free capital mobility, (2) a fixed exchange rate, and (3) an independent monetary policy. Economists call this result the **policy trilemma** (or, more colourfully, the **impossible trinity**). Figure 19-4 illustrates the policy trilemma. A country can choose only two of the three options, which are denoted by the sides of the triangle. In option 1, a country (or monetary union) chooses to have capital mobility and an independent monetary policy but not a fixed exchange rate. Canada, the Eurozone, and the United States have made this choice. Hong Kong and Belize have chosen option 2, in which there is free capital mobility and the exchange rate is fixed, but the country does not have an independent monetary policy. Other countries, like China, have chosen option 3, in which the country has a fixed exchange rate and

Mini-lecture

FIGURE 19-4
The Policy Trilemma
A country (or monetary union) cannot pursue the following three policies at the same time: (1) free capital mobility, (2) a fixed exchange rate, and (3) an independent monetary policy. Instead, it must choose two of the three policies denoted by the sides of the triangle.



pursues an independent monetary policy but does not have free capital mobility because of **capital controls**, or restrictions on the free movement of capital across the borders.

The policy trilemma thus leaves countries with a difficult choice. Do they accept exchange-rate volatility (option 1), give up an independent monetary policy (option 2), or restrict capital flows (option 3)?

APPLICATION How Did China Accumulate US$4 Trillion of International Reserves?

By 2014, China had accumulated US$4 trillion of international reserves (which, however, declined to US$3 trillion by 2020). How did the Chinese get their hands on this vast amount of foreign assets? After all, China is not yet a rich country.

The answer is that in 1994, China pegged its exchange rate to the U.S. dollar at a fixed rate of 12 cents to the yuan (also called the renminbi). China's rapidly growing productivity, accompanied by an inflation rate lower than that of the United States, caused the long-run value of the yuan to increase, leading to a higher relative expected return on yuan assets and a rightward shift of the demand curve for yuan assets. As a result, the Chinese found themselves in the situation depicted in panel (b) of Figure 19-2, a situation in which the yuan was undervalued. To keep the yuan from appreciating above E_{par} to E_1 in the figure, the Chinese central bank has engaged in massive purchases of U.S. dollar assets. Today the Chinese government is one of the largest holders of U.S. government bonds in the world.

The Chinese pegged the yuan at 12 cents from 1994 until 2005. Because the Chinese authorities created substantial roadblocks to capital mobility, they were able to sterilize most of their exchange-rate interventions while maintaining the exchange-rate peg. Nevertheless, their foreign exchange-rate interventions led to a rapidly growing money supply that produced inflationary pressures. As a result, in July 2005, China finally made its peg somewhat more flexible by letting the value of the yuan rise 2.1%, and China subsequently allowed the yuan to appreciate at a gradual pace. The Chinese central bank also indicated that it would no longer fix the value of the yuan to the U.S. dollar but would instead maintain its value relative to a basket of currencies.

After 2014, the Chinese yuan was no longer undervalued, and the Chinese central bank began to purchase yuan, leading to a decline in foreign exchange reserves. By 2020, Chinese international reserves had fallen to US$3 trillion. ◆

Monetary Unions

A variant of a fixed exchange rate regime is a **monetary (or currency) union**, in which a group of countries decides to adopt a common currency, thereby fixing the countries' exchange rates in relation to each other. One of the earliest monetary unions was formed in 1787, when the 13 American colonies formed the United States of America and gave up their individual currencies for the U.S. dollar. The most recently formed monetary union is the European Monetary Union (EMU), which was formed in January 1999, when 11 initial member countries adopted a new joint currency, the euro. There are now 19 countries that have adopted the euro.

The key economic advantage of a monetary union is that it makes trade across borders easier because goods and services in all of the member countries are now priced in the same currency. However, as with any fixed exchange rate regime and free capital mobility, a currency union means that individual countries no longer have their own independent monetary policies with which to address shortfalls of aggregate demand. This disadvantage of a currency union has raised questions about whether the Eurozone will break up, as discussed in the Global box "Will the Euro Survive?"

Managed Float

With the demise of the Bretton Woods system, most exchange rates now change daily in response to market forces, but central banks have not been willing to give up the option of intervening in the foreign exchange market. The ability to prevent large changes in exchange rates makes it easier for firms and individuals purchasing or selling goods abroad to plan for the future. Furthermore, countries with surpluses in their balances of payments frequently do not want to see their currencies appreciate because it makes their goods more expensive abroad and foreign goods cheaper in the home country. Because an appreciation might hurt sales for domestic businesses and increase unemployment, countries with balance-of-payments surpluses often have sold their currencies in the foreign exchange market and acquired international reserves.

Countries with balance-of-payments deficits do not want to see their currencies lose value because it makes foreign goods more expensive for domestic consumers and can stimulate inflation. To keep the value of the domestic currency high, countries with

Global Will the Euro Survive?

The global financial crisis of 2007–2009 led to economic contraction throughout Europe, with the countries in the southern part of the Eurozone hit especially hard. Unemployment in the hard-hit southern countries climbed much faster than in the northern countries such as Germany. Furthermore, with the contraction of their economies, many of the southern countries began to experience large government budget deficits and sovereign debt crises, as described in Chapter 11, in which investors pulled back from purchasing these countries' bonds, sending interest rates to extremely high levels. The resulting collapse of the southern countries' economies could have been alleviated by easier monetary policy that would have stimulated economic activity, but this option was unavailable to these countries because the European Central Bank had to conduct monetary policy for the Eurozone as a whole, which was not suffering as badly as the individual southern countries.

This "straightjacket" effect of the euro has weakened support for the euro in the southern countries, leading to increased talk of abandoning the euro. Support for the euro has also weakened in the stronger northern countries, since they have been called on to provide bailouts to the weaker member countries. Because the stronger countries might wish to abandon their commitment to the euro in order to limit their transfers of funds to the weaker countries, and the weaker countries might wish to abandon their commitment to the euro so that they can boost their economies through more expansionary monetary policy and depreciation of their currencies, there are doubts that the European Monetary Union will survive. However, the euro is still seen by many as an important step in the creation of a more united and powerful Europe, and this political consideration has created strong support for retaining the monetary union.

deficits have often bought their own currencies in the foreign exchange market and given up international reserves.

The current international financial system is a hybrid of a fixed and a flexible exchange rate system. Rates fluctuate in response to market forces but are not determined solely by them. Furthermore, many countries continue to keep the value of their currency fixed against other currencies, as was the case in the European Monetary System before the introduction of the euro, and China from 1994 until 2005.

19.4 CAPITAL CONTROLS

LO 19.4 Summarize the advantages and disadvantages of capital controls.

Emerging market countries are countries that have recently opened up to flows of goods, services, and capital from the rest of the world. Politicians and some economists have advocated that these countries avoid financial instability by restricting capital mobility. Are capital controls a good idea?

Controls on Capital Outflows

Capital outflows can promote financial instability because, when domestic residents and foreigners pull their capital out of a country, the resulting capital outflow forces the country to devalue its currency. Because they want to avoid such instability, politicians in emerging market countries may find capital controls particularly attractive.

Although these controls sound like a good idea, they suffer from several disadvantages. First, empirical evidence indicates that controls on capital outflows are seldom effective during a crisis, because the private sector will find ingenious ways to evade them and will have little difficulty in moving funds out of the country. Second, the evidence suggests that capital flight may actually increase after controls are put into place because confidence in the government is weakened. Third, controls on capital outflows can lead to corruption, since government officials are often paid to look the other way while domestic residents try to move funds abroad. Fourth, controls on capital outflows may lull governments into thinking they do not have to act to reform their financial systems to deal with the crisis, and so opportunities to improve the functioning of the economy are lost.

Controls on Capital Inflows

Although most economists find the arguments against controls on capital outflows persuasive, the idea of controls on capital inflows receives more support. Supporters reason that if speculative capital cannot come in, then it cannot go out suddenly. Thus a potential crisis is averted. Our analysis of financial crises in Chapter 11 provides support for this view by suggesting that capital inflows can lead to credit booms and excessive risk taking on the part of banks, factors that can help trigger a financial crisis.

However, controls on capital inflows have the undesirable feature of possibly blocking funds that could be used for productive investment opportunities from entering a country. Although controls on capital inflows may limit the amount of fuel supplied to lending booms through capital flows, over time they produce substantial distortions and misallocations of resources as households and businesses try to get around them. Indeed, as with controls on capital outflows, controls on capital inflows can lead to corruption. Serious doubts exist over whether capital controls can be effective in today's

environment, in which trade is open and in which the variety of financial instruments makes it easier to get around these controls.

A strong case can be made, however, for improving bank regulation and supervision so that capital inflows are less likely to produce a lending boom that would encourage excessive risk taking by banking institutions. For example, restrictions on the growth of bank borrowing might substantially limit capital inflows. Supervisory controls that focus on the sources of financial fragility, rather than the symptoms, can enhance the efficiency of the financial system rather than hamper it.

19.5 THE ROLE OF THE IMF

LO 19.5 Assess the role of the IMF as an international lender of last resort.

The IMF was originally set up under the Bretton Woods system to help countries deal with balance-of-payments problems and stand by the fixed exchange rates by lending to deficit countries. When the Bretton Woods system of fixed exchange rates collapsed in 1973, the IMF took on new roles.

The IMF continues to function as a data collector and continues to provide technical assistance to its member countries. Although the IMF no longer attempts to encourage fixed exchange rates, its role as an international lender has become more important in recent years. This role first came to the fore in the 1980s during the Third World debt crisis, during which the IMF assisted developing countries in repaying their loans. During the financial crises in Mexico from 1994 to 1995 and in East Asia from 1997 to 1998, the IMF made huge loans to these and other affected countries to help them recover financially and to prevent the spread of the crises to other countries. Then, starting in 2010, the IMF made large loans to Greece, Ireland, and Portugal to help these countries avoid defaults on their government debt. This role of the IMF, in which it acts as an international lender of last resort and helps countries cope with financial instability, is highly controversial.

Should the IMF Act as an International Lender of Last Resort?

We have seen that when a central bank engages in a lender-of-last-resort operation, as the Federal Reserve did during the recent global financial crisis, such an action can minimize the severity of the financial crisis and sometimes even prevent it from happening. When the IMF acts as an international lender of last resort, this role can have similar benefits, particularly when the central bank of the country it is lending to does not have the capacity to cope with the crisis.

However, when the IMF acts as an international lender of last resort, it can lead to serious moral hazard problems. First, knowing that the IMF is willing to take on this role may encourage governments to be profligate, because they know they are likely to be bailed out by the IMF if things get out of hand. Second, governments use IMF funds to protect depositors and other creditors of banking institutions from losses. This safety net creates the moral hazard problem discussed in Chapter 9, because depositors and other creditors have less incentive to monitor these banking institutions and to withdraw their deposits if the institutions take on too much risk.

An international lender of last resort must find ways to limit these two moral hazard problems or it can actually make the situation worse. The international lender of last resort can make it clear that it will extend liquidity only to governments that both

put their fiscal houses in order and implement measures to prevent excessive risk taking by financial institutions. Critics of the IMF, however, believe that the IMF has not put enough pressure on the governments to which it lends to contain these two moral hazard problems, and for this reason they believe the IMF should abandon its role as an international lender of last resort. The IMF has also been criticized for imposing so-called austerity programs that force borrowing governments to cut government spending, raise taxes, and raise interest rates. These austerity programs can be highly contractionary (see Chapter 23) and can lead to high unemployment and even political instability.

The debate over whether the world is better off with the IMF operating as an international lender of last resort is currently a hot one. Much attention is being focused on efforts to make the IMF more effective in performing this role, and redesign of the IMF is at the centre of proposals for a new international financial architecture that will help reduce international financial instability.

19.6 INTERNATIONAL CONSIDERATIONS AND MONETARY POLICY

LO 19.6 Identify the ways in which international monetary policy and exchange-rate arrangements can affect domestic monetary policy operations.

Our analysis in this chapter so far has suggested several ways in which monetary policy can be affected by international matters. Awareness of these effects can have significant implications for the way in which monetary policy is conducted.

Direct Effects of the Foreign Exchange Market on Monetary Policy

When central banks intervene in the foreign exchange market, they acquire or sell off international reserves, and their monetary base is affected. When a central bank intervenes in the foreign exchange market, it gives up some control of its monetary policy. For example, in the early 1970s, the German central bank faced a dilemma. In attempting to keep the German mark from appreciating too much against the U.S. dollar, the Germans acquired huge quantities of international reserves, leading to a high rate of money growth and lower interest rates that the German central bank considered inflationary.

The Bundesbank could have stopped its intervention in the foreign exchange market, thereby reasserting control over its own monetary policy. Such a strategy has a major drawback when a central bank is under pressure not to allow its currency to appreciate: The lower price of imports and higher price of exports that result from an appreciation of its currency will hurt domestic producers and increase unemployment.

Because the U.S. dollar has been a reserve currency, the U.S. monetary base and money supply have not been greatly affected by developments in the foreign exchange market. As long as foreign central banks, rather than the Fed, intervene to keep the value of the dollar from changing, American holdings of international reserves will remain unaffected. Conducting monetary policy is typically easier when a country's currency is a reserve currency.[4]

[4]However, the central bank of a reserve currency country must worry about a shift away from the use of its currency for international reserves.

Exchange Rate Considerations

If a central bank does not want to see its currency fall in value, it may pursue a more contractionary monetary policy to raise the domestic interest rate, thereby strengthening its currency. Similarly, if a country experiences an appreciation in its currency, its domestic industry may suffer from increased foreign competition and may pressure the central bank to ease monetary policy in order to lower the exchange rate.

19.7 TO PEG OR NOT TO PEG: EXCHANGE-RATE TARGETING AS AN ALTERNATIVE MONETARY POLICY STRATEGY

LO 19.7 Summarize the advantages and disadvantages of exchange-rate targeting.

In Chapter 17, we discussed two monetary policy strategies that could be used to promote price stability, one of which is inflation targeting. Another strategy that uses a strong nominal anchor to promote price stability is called **exchange-rate targeting** (sometimes referred to as an **exchange-rate peg**).

Targeting the exchange rate is a monetary policy strategy with a long history. It can take the form of fixing the value of the domestic currency to a commodity such as gold, the key feature of the gold standard described earlier in the chapter. More recently, fixed exchange rate regimes have involved fixing the value of the domestic currency to that of a large, low-inflation country like the United States (the *anchor country*). Another alternative is to adopt a *crawling target* or *peg*, in which a currency is allowed to depreciate at a steady rate so that the inflation rate of the pegging country can be higher than that of the anchor country.

Advantages of Exchange-Rate Targeting

Exchange-rate targeting has several advantages. First, the nominal anchor of an exchange-rate target directly contributes to the goal of keeping inflation under control by tying the inflation rate for internationally traded goods to the rate in the anchor country. The exchange-rate target is able to do this because the foreign price of internationally traded goods is set by the world market, whereas the domestic price of these goods is fixed by the exchange-rate target. For example, until 2002, the exchange rate for the Argentine peso was exactly one to the dollar, so a bushel of wheat traded internationally at five dollars had its price set at five pesos. If the exchange-rate target is credible (i.e., expected to be adhered to), the exchange-rate target has the added benefit of anchoring inflation expectations to the inflation rate in the anchor country.

Second, an exchange-rate target provides an automatic rule for the conduct of monetary policy that helps mitigate the time-inconsistency problem described in Chapter 17. As we saw earlier in the chapter, an exchange-rate target forces a tightening of monetary policy when there is a tendency for the domestic currency to depreciate and a loosening of policy when there is a tendency for the domestic currency to appreciate, so discretionary monetary policy is less of an option. The central bank will therefore be constrained from falling into the time-inconsistency trap of trying to expand output and employment in the short run by pursuing overly expansionary monetary policy.

Third, an exchange-rate target has the advantage of simplicity and clarity, which makes it easily understood by the public. A "sound currency" is an easy-to-understand rallying cry for monetary policy. In the past, for example, this aspect of simplicity was important in France, where an appeal to the "franc fort" (strong franc) was often used to justify tight monetary policy.

Given its advantages, it is not surprising that exchange-rate targeting has been used successfully to control inflation in industrialized countries. Both France and the United Kingdom, for example, successfully used exchange-rate targeting to lower inflation by tying the values of their currencies to the German mark. In 1987, when France first pegged its exchange rate to the mark, its inflation rate was 3%, two percentage points above the German inflation rate. By 1992, its inflation rate had fallen to 2%, a level that can be argued is consistent with price stability and that was actually below the German inflation rate. By 1996, the French and German inflation rates had converged to a number slightly below 2%. Similarly, after pegging its currency to the German mark in 1990, the United Kingdom was able to lower its inflation rate from 10% to 3% by 1992, when it was forced to abandon the ERM.

Exchange-rate targeting has also been an effective means of reducing inflation quickly in emerging market countries. For example, before the devaluation in Mexico in 1994, its exchange-rate target enabled it to bring inflation down from levels above 100% in 1988 to levels below 10% in 1994.

Disadvantages of Exchange-Rate Targeting

Despite the inherent advantages of exchange-rate targeting, several serious criticisms of this strategy can be made. The problem (as we saw earlier in the chapter) is that with capital mobility, the targeting country can no longer pursue its own independent monetary policy and use it to respond to domestic shocks that are independent of those hitting the anchor country. Furthermore, an exchange-rate target means that shocks to the anchor country are directly transmitted to the targeting country, because changes in interest rates in the anchor country lead to corresponding changes in interest rates in the targeting country.

A striking example of these weaknesses occurred when Germany was reunified in 1990. In response to concerns about inflationary pressures arising from reunification and the massive fiscal expansion required to rebuild East Germany, long-term German interest rates rose until February 1991, and short-term rates rose until December 1991. This shock to the anchor country in the ERM was transmitted directly to the other countries in the ERM whose currencies were pegged to the mark, and their interest rates rose in tandem with the German rates. Continued adherence to the exchange-rate target slowed economic growth and increased unemployment in countries such as France that remained in the ERM exchange-rate peg.

A second problem with exchange-rate targets is that they leave countries open to speculative attacks on their currencies. Indeed, one aftermath of German reunification was the foreign exchange crisis of September 1992. As we saw earlier, the tight monetary policy practised in Germany following reunification subjected the countries of the ERM to a negative demand shock that led to declines in economic growth and increases in unemployment. It was certainly feasible for the governments of these countries to keep their exchange rates fixed relative to the mark under these circumstances, but speculators began to question whether these

countries' commitment to the exchange-rate peg would eventually weaken. Speculators reasoned that these countries would not tolerate the rise in unemployment that resulted from keeping interest rates high enough to fend off attacks on their currencies.

At this stage, speculators were, in effect, presented with a one-way bet, because the currencies of countries like France, Spain, Sweden, Italy, and the United Kingdom could go in only one direction and depreciate against the mark. Selling these currencies before the likely depreciation occurred gave speculators an attractive profit opportunity with potentially high expected returns. The result was the speculative attack of September 1992. Only in France was the commitment to the fixed exchange rate strong enough that France did not devalue. The governments of the other ERM countries were unwilling to defend their currencies at all costs and eventually allowed their currencies to fall in value.

The different responses of France and the United Kingdom after the September 1992 exchange-rate crisis illustrate the potential cost of an exchange-rate target. On the one hand, France, which continued to peg its currency to the mark and thus was unable to use monetary policy to respond to domestic conditions, found that economic growth remained slow after 1992, and unemployment increased. The United Kingdom, on the other hand, which dropped out of the ERM exchange-rate peg and adopted inflation targeting, enjoyed much better economic results: Economic growth was higher, the unemployment rate fell, and yet the inflation rate was not much worse than France's.

In contrast to industrialized countries, emerging market countries (including the transition countries of Eastern Europe) may not lose all that much by giving up an independent monetary policy when they target exchange rates. Because many emerging market countries have not developed the political or monetary institutions that are necessary for the successful use of discretionary monetary policy, they may have little to gain but a lot to lose from an independent monetary policy. Thus they are better off if they adopt, in effect, the monetary policy of a country like the United States through targeting exchange rates than if they pursue their own independent policy. This is one of the reasons so many emerging market countries have adopted exchange-rate targeting.

Nonetheless, exchange-rate targeting is highly dangerous for these countries because it leaves them open to speculative attacks that can have far more serious consequences for their economies than for those of industrialized countries. Indeed, the successful speculative attacks on Mexico in 1994, East Asia in 1997, and Argentina in 2002 plunged their economies into full-scale financial crises that devastated their countries.

An additional disadvantage of an exchange-rate target is that it can weaken the accountability of policymakers, particularly in emerging market countries. Because exchange-rate targeting fixes the exchange rate, it eliminates an important signal that can help constrain monetary policy from becoming too expansionary and thereby limit the time-inconsistency problem. In industrialized countries, particularly in Canada and the United States, the bond market provides an important signal about the stance of monetary policy. Overly expansionary monetary policy or strong political pressure to engage in overly expansionary monetary policy produces an inflation scare, in which inflation expectations surge, interest rates rise (because of the Fisher effect, described in Chapter 5), and long-term bond prices decline sharply. Because both central banks and politicians want to avoid this kind of scenario, overly expansionary monetary policy is less likely.

In many countries, particularly emerging market countries, the long-term bond market is essentially nonexistent. Under a floating exchange rate regime, however, if monetary policy is too expansionary, the exchange rate will depreciate. In these countries, the daily fluctuations of the exchange rate can, like the bond market in Canada, provide an early warning signal that monetary policy is too expansionary. Just as the fear of a visible inflation scare in the bond market constrains central bankers from pursuing overly expansionary monetary policy and constrains politicians from putting pressure on the central bank to do so, fear of exchange-rate depreciations can make overly expansionary monetary policy, and therefore the time-inconsistency problem, less likely.

The need for signals from the foreign exchange market may be even more acute for emerging market countries because the balance sheets and actions of their central banks are not as transparent as those of industrialized countries. Targeting the exchange rate can make it even harder to ascertain the central bank's policy actions. The public is less able to keep a watch on the central bank and the politicians pressuring it, which makes it easier for monetary policy to become too expansionary.

When Is Exchange-Rate Targeting Desirable for Industrialized Countries?

Given the above disadvantages of exchange-rate targeting, when might it make sense? In industrialized countries, the biggest cost to exchange-rate targeting is the loss of an independent monetary policy to deal with domestic considerations. If an independent, domestic monetary policy can be conducted responsibly, the loss of such a policy can be a serious cost indeed, as the comparison between the post-1992 experiences of France and the United Kingdom indicates. However, not all industrialized countries are capable of conducting their own monetary policies successfully, either because the central bank is not independent or because political pressures on the central bank lead to an inflationary bias in monetary policy. In these cases, giving up independent control of domestic monetary policy may not be such a great loss, while the advantages of having monetary policy determined by a better-performing central bank in the anchor country can be substantial.

Italy provides a good example. It is not a coincidence that the Italian public had the most favourable attitude of all the European countries toward the European Monetary Union. The past record of Italian monetary policy was not good, and the Italian public recognized that a monetary policy controlled by more responsible outsiders had benefits that far outweighed the costs to the country of losing the ability to focus monetary policy on domestic considerations.

A second reason why industrialized countries might find exchange-rate targeting useful is that it encourages integration of the domestic economy with the economies of neighbouring countries. Clearly, this was the rationale for the long-standing pegging of the exchange rate to the deutsche mark by countries such as Austria and the Netherlands and the more recent exchange-rate pegs that preceded the formation of the European Monetary Union.

To sum, exchange-rate targeting for industrialized countries is probably not the best monetary policy strategy for controlling the overall economy unless (1) domestic monetary and political institutions are not conducive to good monetary policymaking or (2) an exchange-rate target has other important benefits that have nothing to do with monetary policy.

When Is Exchange-Rate Targeting Desirable for Emerging Market Countries?

In countries with particularly weak political and monetary institutions that have experienced continued bouts of hyperinflation, a characterization that applies to many emerging market (including transition) countries, exchange-rate targeting may be the only way to break inflationary psychology and stabilize the economy. In this situation, exchange-rate targeting is the stabilization policy of last resort. However, if the exchange-rate targeting regimes in emerging market countries are not transparent, they are more likely to break down, often resulting in disastrous financial crises.

Are there exchange-rate strategies that will make it less likely that the exchange rate regime will break down in an emerging market country? Two such strategies, currency boards and dollarization, have received increased attention in recent years.

Currency Boards

A **currency board** is an arrangement in which the domestic currency is backed 100% by a foreign currency (say, U.S. dollars), and the note-issuing authority, whether it be the central bank or the government, establishes a fixed exchange rate to this foreign currency and stands ready to exchange domestic currency for the foreign currency at this rate whenever the public so requests. A currency board is just a variant of a fixed exchange-rate target in which the commitment to the fixed exchange rate is especially strong because the conduct of monetary policy is, in effect, put on autopilot and taken completely out of the hands of the central bank and the government. In contrast, the typical fixed or pegged exchange rate regime does allow the monetary authorities some discretion in their conduct of monetary policy because they can still adjust interest rates or print money.

A currency board arrangement thus has important advantages over a monetary policy strategy that just uses an exchange-rate target. First, the money supply can expand only when foreign currency is exchanged for domestic currency at the central bank. Therefore, the increased amount of domestic currency is matched by an equal increase in foreign exchange reserves. The central bank no longer has the ability to print money and thereby cause inflation. Second, the currency board setup involves a stronger commitment by the central bank to the fixed exchange rate and may therefore be effective in bringing inflation down quickly and in decreasing the likelihood of a successful speculative attack against the currency.

Although currency boards solve the transparency and commitment problems inherent in an exchange-rate-target regime, they suffer from some of the same shortcomings: the loss of an independent monetary policy, increased exposure of the economy to shocks from the anchor country, and the loss of the central bank's ability to create money and act as a lender of last resort. Other means must therefore be used to cope with potential banking crises. In addition, if a speculative attack on a currency board occurs, the exchange of domestic currency for foreign currency leads to a sharp contraction of the money supply, which can be highly damaging to the economy.

Currency boards have been established in the territory of Hong Kong (1983) and in countries such as Argentina (1991), Estonia (1992), Lithuania (1994), Bulgaria (1997), and Bosnia (1998). Argentina's currency board, which operated from 1991 until 2002 and required the central bank to exchange U.S. dollars for new pesos at a fixed exchange rate of one to one, is one of the most interesting. For more on this subject, see the Global box "Argentina's Currency Board."

Global Argentina's Currency Board

Argentina has a long history of monetary instability, with inflation rates fluctuating dramatically over the years, sometimes surging to beyond 1000% per year. To end this cycle of inflationary surges, Argentina decided to adopt a currency board in April 1991. The Argentine currency board worked as follows. Under Argentina's convertibility law, the peso/U.S. dollar exchange rate was fixed at one to one, and a member of the public could go to the Argentine central bank and exchange a peso for a U.S. dollar, or vice versa, at any time.

The early years of Argentina's currency board looked stunningly successful. Inflation, which was running at an 800% annual rate in 1990, fell to less than 5% by the end of 1994, and economic growth was rapid, averaging almost 8% per year from 1991 to 1994. In the aftermath of the Mexican peso crisis, however, concern about the health of the Argentine economy led the public to pull money out of the banks (deposits fell by 18%) and exchange pesos for U.S. dollars, thus causing a contraction of the Argentine money supply. The result was a sharp drop in Argentine economic activity, with real GDP shrinking by more than 5% in 1995 and the unemployment rate jumping above 15%. Only in 1996 did the economy begin to recover.

Because the central bank of Argentina had no control over monetary policy under the currency board system, it was virtually helpless to counteract the contractionary monetary effects stemming from the public's behaviour. Furthermore, because the currency board did not allow the central bank to create pesos and lend them to the banks, it had very little capability to act as a lender of last resort. With help from international agencies such as the IMF, the World Bank, and the Inter-American Development Bank, which lent Argentina more than US$5 billion in 1995 to help shore up its banking system, the currency board survived.

In 1998, however, Argentina entered another recession, one that was both severe and very long-lasting. By the end of 2001, unemployment had reached nearly 20%, a level comparable to that experienced in the United States during the Great Depression of the 1930s. The result was civil unrest and the fall of the elected government as well as a major banking crisis and a default on nearly US$150 billion of government debt. Because the Central Bank of Argentina had no control over monetary policy under the currency board system, it was unable to use monetary policy to expand the economy and get Argentina out of its recession. In addition, because the currency board limited the central bank's ability to create pesos and act as a lender of last resort, the central bank was unable to help banks weather the banking crisis. In January 2002, the currency board finally collapsed, the peso depreciated by more than 70%, and inflation shot through the roof. The result was a full-scale financial crisis that led to an extremely severe depression. Clearly, the Argentine public is not as enamoured of its currency board as it once was.*

*The Argentine financial crisis is described in detail in an additional chapter, "Financial Crises in Emerging Market Economies," which can be found online.

Dollarization

Another solution to the transparency problem and doubts about commitment to the exchange-rate target is **dollarization**, the adoption of a sound currency, like the euro or the U.S. dollar, as a country's money. Indeed, dollarization is just another variant of a fixed exchange-rate target but with an even stronger commitment mechanism than that provided by a currency board. A currency board can be abandoned, allowing a change in the value of the currency, but a change of value is impossible with dollarization: A dollar bill is always worth one dollar, whether it is held in the United States or outside it.

Dollarization has been advocated as a monetary policy strategy for emerging market countries: It was discussed actively by Argentine officials in the aftermath of the devaluation of the Brazilian real in January 1999 and was adopted by Ecuador in March 2000. Dollarization's key advantage is that it completely eliminates the possibility of a speculative attack on the domestic currency (because there is no such currency). (Such an attack is still a danger under a currency board arrangement.)

Dollarization is subject to the usual weaknesses of an exchange-rate target (the loss of an independent monetary policy, increased exposure of the economy to shocks from the anchor country, and the inability of the central bank to create money and act as a lender of last resort). Dollarization has one additional weakness not characteristic of currency boards or other exchange-rate-target regimes. Because a country adopting dollarization no longer has its own currency, it loses the revenue that a government receives by issuing money, which is called **seignorage**. Because governments (or their central banks) do not have to pay interest on their domestic currency, they earn revenue (seignorage) by using this currency to purchase income-earning assets such as bonds. In the case of the Federal Reserve in the United States, this revenue is usually in excess of $30 billion per year. If an emerging market country dollarizes and gives up its currency, it needs to make up this loss of revenue somehow, and this is not always easy for a poor country.

SUMMARY

1. An unsterilized central bank intervention in which the domestic currency is sold to purchase foreign assets leads to a gain in international reserves, an increase in the money supply, and a depreciation of the domestic currency. Available evidence suggests, however, that sterilized central bank interventions have little long-term effect on the exchange rate.
2. The balance of payments is a bookkeeping system used to record all transactions between a country and foreign countries. The current account balance is the sum of the trade balance, net investment income, and transfers. A current account deficit is financed by the financial account balance, and so a current account deficit indicates that a country is becoming more indebted to foreigners.
3. Before World War I, the gold standard was predominant. Currencies were convertible into gold, thus fixing exchange rates between them. After World War II, the Bretton Woods system and the IMF were established to promote a fixed exchange rate system in which the U.S. dollar, the reserve currency, was convertible into gold. The Bretton Woods system collapsed in 1971. We now have an international financial system that has elements of a managed float and a fixed exchange rate system. Some exchange rates fluctuate from day to day, even though central banks intervene in the foreign exchange market, while other exchange rates are fixed.
4. Controls on capital outflows receive support because they may prevent domestic residents and foreigners from pulling capital out of a country during a crisis and thereby make devaluation less likely. Controls on capital inflows make sense under the theory that, if speculative capital cannot flow in, then it also cannot go out suddenly. Thus a crisis is averted. However, capital controls suffer from several disadvantages: They are seldom effective, they often lead to corruption, and they may encourage governments to avoid taking the steps that are necessary to reform their financial systems to deal with the crisis.
5. The IMF's role as an international lender of last resort can help lessen the severity of financial crises. However, the IMF's role as an international lender of last resort also creates serious moral hazard problems because it encourages governments to be profligate and also encourages excessive risk taking by financial institutions, making a financial crisis more likely.
6. Two international considerations affect the conduct of monetary policy: direct effects of the foreign

exchange market on monetary policy and exchange rate considerations. Inasmuch as the United States has been a reserve currency country in the post–World War II period, U.S. monetary policy has been less affected than that of other countries by developments in the foreign exchange market. However, in recent years, exchange rate considerations have played a more prominent role in influencing U.S. monetary policy.

7. Exchange-rate targeting has the following advantages as a monetary policy strategy: (1) It directly keeps inflation under control by tying the inflation rate for internationally traded goods to that found in the anchor country to which the targeting country's currency is pegged; (2) it provides an automatic rule for the conduct of monetary policy that helps mitigate the time-inconsistency problem; and (3) it is simple and clear. Exchange-rate targeting also has some serious disadvantages: (1) It results in a loss of independent monetary policy in the pegging country; (2) it leaves the pegging country open to speculative attacks; and (3) it can weaken the accountability of policymakers because the exchange-rate signal is lost. Two strategies that make it less likely that exchange-rate targeting will break down are currency boards, in which the central bank stands ready to automatically exchange domestic currency for foreign currency at a fixed rate, and dollarization, in which a sound currency like the euro or the U.S. dollar is adopted as a country's money.

MyLab Economics Make the grade with MyLab Economics: Select end-of-chapter questions can be found on MyLab Economics. You can practise them as often as you want, and most feature step-by-step guided instructions to help you find the right answer.

QUESTIONS

1. If the Bank of Canada sells dollars in the foreign exchange market but conducts an offsetting open market operation to sterilize the intervention, what will be the effect on international reserves, the money supply, and the exchange rate?
2. If the Bank of Canada buys dollars in the foreign exchange market but does not sterilize the intervention, what will be the impact on international reserves, the money supply, and the exchange rate?
3. For each of the following, identify whether they increase or decrease the current account balance:
   a. A Canadian citizen's purchase of an airline ticket from Air France
   b. A Japanese citizen's purchase of California oranges
   c. $50 million of foreign aid to Honduras
   d. A worker in Toronto who sends money to their parents in Mexico
   e. The services a Canadian accounting firm provides to a German firm
4. Suppose that you travel to Cali (Colombia), where the exchange rate is 1 Canadian dollar to 2900 Colombian pesos. As you enter a McDonald's restaurant, you realize you need 17 400 Colombian pesos to buy a Big Mac. Assuming a Big Mac sells for $5 in Canada, would you say that the Colombian peso is over- or undervalued in terms of PPP?
5. Refer to the previous exercise. Which type of foreign market intervention must the central bank of Colombia conduct to keep the exchange rate at a level where the currency is not under- or overvalued in terms of PPP?
6. What would be the effect of a devaluation on a country's imports and exports? If a country imports most of the goods included in the basket of goods and services used to calculate the CPI, what do you think the effect will be on this country's inflation rate?
7. Under the gold standard, if Britain became more productive relative to Canada, what would happen to the money supply in the two countries? Why would the changes in the money supply help preserve a fixed exchange rate between Canada and Britain?
8. What is the exchange rate between dollars and Swiss francs if one dollar is convertible into 1/20 ounce of gold, and one Swiss franc is convertible into 1/40 ounce of gold?

9. "Inflation is not possible under the gold standard." Is this statement true, false, or uncertain? Explain your answer.
10. What are some of the disadvantages of China's pegging the yuan to the U.S. dollar?
11. If a country's par exchange rate was undervalued during the Bretton Woods fixed exchange rate regime, what kind of intervention would that country's central bank be forced to undertake, and what effect would the intervention have on the country's international reserves and money supply?
12. Why might a country that is suffering a recession not want to intervene in the foreign exchange market if its currency is overvalued? Assume this country participates in a fixed exchange rate regime.
13. "If a country wants to keep its exchange rate from changing, it must give up some control over its monetary policy." Is this statement true, false, or uncertain? Explain your answer.
14. Why is it that in a pure, flexible exchange rate system, the foreign exchange market has no direct effect on the money supply? Does this mean that the foreign exchange market has no effect on monetary policy?
15. Why did the exchange-rate peg lead to difficulties for the countries in the ERM after the German reunification?
16. How can exchange-rate targets lead to a speculative attack on a currency?
17. The IMF does not enjoy a great reputation in many countries that were recipients of IMF loans or bailouts. Explain why many citizens were not happy with the role played by the IMF.
18. How can the long-term bond market help reduce the time-inconsistency problem for monetary policy? Can the foreign exchange market also perform this role?
19. What are the key advantages of exchange-rate targeting as a monetary policy strategy?
20. When is exchange-rate targeting likely to be a sensible strategy for industrialized countries? When is exchange-rate targeting likely to be a sensible strategy for emerging market countries?

APPLIED PROBLEMS

21. Calculate the overvaluation of the Thai baht (THB) if you can get 34.6 THB per Canadian dollar at the exchange counter, but a lunch menu that costs 25 Canadian dollars in Montreal sells for 948.25 THB in Bangkok.
22. Suppose the Bank of Canada purchases $1 000 000 worth of foreign assets.
    a. If the Bank of Canada purchases the foreign assets with $1 000 000 in currency, show the effect of this open market operation using T-accounts. What happens to the monetary base?
    b. If the Bank of Canada purchases the foreign assets by selling $1 000 000 in T-bills, show the effect of this open market operation using T-accounts. What happens to the monetary base?
23. Suppose the Mexican central bank chooses to peg the peso to the U.S. dollar and commits to a fixed peso/dollar exchange rate. Use a graph of the market for peso assets (foreign exchange) to show and explain how the peg must be maintained if a shock in the U.S. economy forces the Fed to pursue contractionary monetary policy. What does this say about the ability of central banks to address domestic economic problems while maintaining a pegged exchange rate?

DATA ANALYSIS PROBLEMS

1. Visit the Bank of Canada's website at www.bankofcanada.ca. Click on Statistics and then Exchange Rates.
    a. What is the Canadian effective exchange rate index (CEER)?
    b. How many foreign currencies are included in the index?
2. **Real-Time Data Analysis** Go to the St. Louis Federal Reserve FRED database and find U.S. data on net exports (NETEXP), transfers (A123RC1Q027SBEA), and the current account balance (NETFI).
    a. Calculate net investment income for the most recent quarter available and for the same quarter a year earlier.

b. Calculate the percentage change in the current account balance from the same quarter one year earlier. Which one of the three items making up the current account balance had the largest effect in percentage terms on the change of the current account? Which one had the smallest effect?

3. **Real-Time Data Analysis** Go to the St. Louis Federal Reserve FRED database and find data on the monthly U.S. dollar exchange rate to the Chinese yuan (EXCHUS), the Canadian dollar (EXCAUS), and the South Korean won (EXKOUS). Download the data into a spreadsheet.
   a. For the most recent five-year period of data available, use the average, max, min, and stdev functions in Excel to calculate the average, highest, and lowest exchange-rate values as well as the standard deviation of the exchange rate to the dollar (this is an absolute measure of the volatility of the exchange rate).
   b. Using the maximum and minimum values of each exchange rate over the past five years, calculate the ratio of the difference between the maximum and minimum values to the average level of the exchange rate (expressed as a percentage by multiplying by 100). This value gives an indication of how tightly the exchange rate moves. Based on your results, which of the three countries is most likely to peg its currency to the U.S. dollar? How does this country's currency compare with the other two?
   c. Calculate the ratio of the standard deviation to the average exchange rate over the past five years (expressed as a percentage by multiplying by 100). This value gives an indication of how volatile the exchange rate is. Based on your results, which of the three currencies is most likely to be pegged to the U.S. dollar? How does this currency compare with the other two?

PART 7

Monetary Theory

Crisis and Response: The Great Recession and the Coronavirus Recession

In 2007 and 2008, the Canadian economy was hit by a perfect storm of formidable shocks. By the end of 2007, oil prices had risen from US$60 per barrel at the beginning of the year to $100, reaching a peak of over $140 per barrel in July 2008. The oil price shock was both contractionary and inflationary and as a result led to both higher inflation and higher unemployment—not to mention many unhappy drivers at the gas pumps. If this supply shock were not bad enough, the global financial crisis hit the economy starting in August 2007 and caused a contraction in both household and business spending. This shock led to a further rise in unemployment, with some weakening of inflationary pressure further down the road.

The result of this perfect storm of adverse shocks was the most severe economic contraction since the Great Depression, now named the Great Recession, with unemployment rising from the 6% level in 2006 and 2007 to over 10% by the end of 2009. Inflation also accelerated from 2.6% in 2007 to over 5% by the middle of 2008, but with the increase in the unemployment rate and the decline of oil and other commodity prices by the fall of 2008, inflation rapidly came back down again.

The Canadian economy was hit by an even larger adverse shock in March 2020, when the coronavirus pandemic led to massive lockdowns. Not only was there a dramatic cut in production but household and business spending collapsed as well. The result was that from February 2020 to May 2020, the unemployment rate rose from 5.6% to over 14%, a far higher level than it reached in 2009. With the collapse of oil prices and the rise in unemployment, inflation fell from above the 2% level before the pandemic to less than −0.4% in May.

Both recessions led to aggressive monetary policy actions to address the contractionary forces in the economy, with the Bank of Canada cutting its policy rate, the overnight interest rate, to zero while also creating new lending programs to shore up both financial institutions and businesses. The federal government also took actions to stabilize the economy, especially during the coronavirus pandemic. It implemented a massive fiscal stimulus package of close to a quarter of a trillion dollars in 2020. Although this stimulus package helped boost GDP, however, its was overwhelmed by the negative shocks from coronavirus pandemic, so the Canadian economy went into a tailspin.

The impact of adverse shocks highlights the need to understand how monetary and other government policies affect inflation and economic activity. Chapter 20 discusses how the quantity theory of money explains inflation in the long run and how theories of the demand for money have evolved. Aggregate supply and demand analysis, the basic framework that will enable us to study the effects of monetary policy on output and inflation, is then developed in Chapters 21–24. In Chapter 25, we expand on aggregate supply and demand analysis to understand how monetary policy can be used to stabilize the economy and inflation. Chapter 26 outlines the transmission mechanisms through which monetary policy affects the aggregate economy.

20 Quantity Theory, Inflation, and the Demand for Money

Learning Objectives

20.1 Assess the relationship between money growth and inflation in the short run and the long run, as implied by the quantity theory of money.

20.2 Identify the circumstances under which budget deficits can lead to inflationary monetary policy.

20.3 Summarize the three motives underlying the liquidity preference theory of money demand.

20.4 Identify the factors underlying the portfolio choice theory of money demand.

20.5 Assess and interpret the empirical evidence on the validity of the liquidity preference and portfolio theories of money demand.

Preview

In earlier chapters, we spent a lot of time and effort learning what the money supply is, how it is determined, and what part the Bank of Canada plays in it. Now we are ready to explore the role of the money supply and monetary policy in determining inflation and total production of goods and services (aggregate output) in the economy. The study of the effects of money and monetary policy on the economy is called **monetary theory**, and we examine this branch of economics in the chapters of Part 7.

When economists mention *supply*, the word *demand* is sure to follow, and the discussion of money is no exception. Understanding the supply of money is an essential building block in understanding how monetary policy affects the economy, because understanding the supply of money suggests the factors that influence the quantity of money in the economy. Not surprisingly, another essential part of monetary theory is the demand for money.

After discussing the quantity theory of money and its link to the demand for money, we delve more deeply into the factors that determine the demand for money. A central question in monetary theory is whether or to what extent the quantity of money demanded is affected by changes in interest rates. Because this issue is crucial to how we view money's effects on aggregate economic activity, we will focus on the role of interest rates in the demand for money.

20.1 QUANTITY THEORY OF MONEY

LO 20.1 Assess the relationship between money growth and inflation in the short run and the long run, as implied by the quantity theory of money.

Developed by the classical economists in the nineteenth and early twentieth centuries, the quantity theory of money explains how the nominal value of aggregate income is determined. Because the theory also tells us how much money is held for a given amount of aggregate income, it is a theory of the demand for money. The most important feature of this theory is that it suggests that interest rates have no effect on the demand for money.

Velocity of Money and Equation of Exchange

The clearest exposition of the classical quantity theory approach is found in the work of the American economist Irving Fisher in his influential book *The Purchasing Power of Money*, published in 1911. Fisher wanted to examine the link between the total

quantity of money M (the money supply) and the total amount of spending on final goods and services produced in the economy $P \times Y$, where P is the price level and Y is aggregate output (income). (Total spending $P \times Y$ is also thought of as aggregate nominal income for the economy or as nominal GDP.) The concept that provides the link between M and $P \times Y$ is called the **velocity of money** (often abbreviated to *velocity*), the average number of times per year (turnover) that a dollar is spent in buying the total amount of goods and services produced in the economy. Velocity V is defined more precisely as total spending $P \times Y$ divided by the quantity of money M:

$$V = \frac{P \times Y}{M} \tag{1}$$

If, for example, nominal GDP $(P \times Y)$ in a year is \$10 trillion and the quantity of money (M) is \$2 trillion, we can calculate velocity as follows:

$$V = \frac{\$10 \text{ trillion}}{\$2 \text{ trillion}} = 5$$

The value of 5 for velocity means that the average dollar is spent five times in purchasing final goods and services in the economy.

By multiplying both sides of Equation 1 by M, we obtain the **equation of exchange**, which relates nominal income to the quantity of money and velocity:

$$M \times V = P \times Y \tag{2}$$

The equation of exchange thus states that the quantity of money multiplied by the number of times this money is spent in a given year must equal nominal income (the total nominal amount spent on goods and services in that year).[1]

As it stands, Equation 2 is nothing more than an identity—a relationship that is true by definition. It does not tell us, for instance, that when the money supply M changes, nominal income $(P \times Y)$ changes in the same direction; a rise in M, for example, might be offset by a fall in V, leaving $M \times V$ (and therefore $P \times Y$) unchanged. To convert the equation of exchange (an *identity*) into a *theory* of how nominal income is determined, we must first understand the factors that determine velocity.

Determinants of Velocity Irving Fisher reasoned that velocity is determined by the institutions within an economy that affect the ways in which individuals conduct transactions. If people use charge accounts and credit cards to conduct their transactions, as they often do today, and consequently use money less often when making purchases, less money is required to conduct the transactions generated by nominal income (M falls relative to $P \times Y$), and so velocity $(P \times Y)/M$ increases. Conversely, if it is more convenient for purchases to be paid for with cash, cheques, or debit cards

[1]Fisher actually first formulated the equation of exchange in terms of the nominal value of transactions in the economy PT:

$$MV_T = PT$$

where
P = average price per transaction
T = number of transactions conducted in a year
$V_T = PT/M$ = transactions velocity of money

Because the nominal value of transactions T is difficult to measure, the quantity theory has been formulated in terms of aggregate output Y as follows: T is assumed to be proportional to Y so that $T = vY$, where v is a constant of proportionality. Substituting vY for T in Fisher's equation of exchange yields $MV_T = vPY$, which can be written as Equation 2 in the text, in which $V = V_T/v$.

(all of which are money), more money is used to conduct the transactions generated by the same level of nominal income, and so velocity falls. Fisher took the view that the institutional and technological features of the economy would affect velocity only slowly over time, so velocity would normally be reasonably constant in the short run.

Demand for Money Fisher's quantity theory can also be interpreted in terms of the **demand for money**, the quantity of money that people want to hold.

Because the quantity theory of money tells us how much money is held for a given amount of nominal spending, it is, in fact, a theory of the demand for money. To illustrate, let's first divide both sides of the equation of exchange (Equation 2) by V to yield the following:

$$M = \frac{1}{V} \times PY$$

When the money market is in equilibrium, money supply equals money demand, so we can replace M in the equation with M^d. In addition, since velocity is assumed to be constant in the quantity theory of money, we can replace $1/V$ with a constant k. Substituting k for $1/V$ and M^d for M, we can rewrite the equation as

$$M^d = k \times PY \tag{3}$$

Equation 3 tells us that because k is constant, the level of transactions generated by a fixed level of nominal income PY determines the quantity of money M^d that people will demand. Therefore, Fisher's theory suggests that the demand for money is purely a function of income, and interest rates have no effect on the demand for money.[2]

From the Equation of Exchange to the Quantity Theory of Money

Fisher's view that velocity is fairly constant in the short run, so that $V = \overline{V}$, transforms the equation of exchange into the **quantity theory of money**, which states that nominal income (spending) is determined solely by movements in the quantity of money M.

$$P \times Y = M \times \overline{V} \tag{4}$$

The quantity theory equation above indicates that when the quantity of money M doubles, $M \times \overline{V}$ doubles, and so must $P \times Y$, the value of nominal income. To illustrate, let's assume that velocity is 5, nominal income (GDP) is initially \$10 trillion, and the money supply is \$2 trillion. If the money supply doubles to \$4 trillion, the quantity theory of money suggests that nominal income will double to \$20 trillion $(= 5 \times \$4 \text{ trillion})$.

Quantity Theory and the Price Level

Because the classical economists (including Fisher) thought that wages and prices were completely flexible, they believed that the level of aggregate output Y produced in the economy during normal times would remain at the full-employment level. Hence Y in the equation of exchange could be treated as reasonably constant in the short run and

[2]While Fisher was developing his quantity theory approach to the demand for money, a group of classical economists in Cambridge, England, led by Alfred Marshall and A. C. Pigou, came to similar conclusions, although via slightly different reasoning. They derived Equation 3 by recognizing that two properties of money motivate people to hold it: its utility as a medium of exchange and its value as a store of wealth.

thus could be assigned a fixed value of $\overline{Y}$ in Equation 4. Dividing both sides of Equation 4 by $\overline{Y}$, we can write the price level as follows:

$$P = \frac{M \times \overline{V}}{\overline{Y}} \tag{5}$$

The quantity theory of money as represented by Equation 5 implies that if M doubles, P must also double in the short run, because $\overline{V}$ and $\overline{Y}$ are constant. In our example, if aggregate output is \$10 trillion, velocity is 5, and the money supply is \$2 trillion, then the price level equals 1.0.

$$P = \frac{\$2 \text{ trillion} \times 5}{\$10 \text{ trillion}} = \frac{\$10 \text{ trillion}}{\$10 \text{ trillion}} = 1.0$$

When the money supply doubles to \$4 trillion, the price level must also double, to 2.0, because

$$P = \frac{\$4 \text{ trillion} \times 5}{\$10 \text{ trillion}} = \frac{\$20 \text{ trillion}}{\$10 \text{ trillion}} = 2.0$$

Classical economists relied on the quantity theory of money to explain movements in the price level. In their view, ***changes in the quantity of money lead to proportional changes in the price level.***

Quantity Theory and Inflation

We now transform the quantity theory of money into a theory of inflation. You might recall from high school the mathematical fact that the percentage change $(\%\Delta)$ of a product of two variables is approximately equal to the sum of the percentage changes of the individual variables. In other words,

$$\begin{aligned}\text{percentage change in } (x \times y) = &\ (\text{percentage change in } x) \\ &+ (\text{percentage change in } y)\end{aligned}$$

Using this mathematical fact, we can rewrite the equation of exchange as follows:

$$\%\Delta M + \%\Delta V = \%\Delta P + \%\Delta Y$$

Subtracting $\%\Delta Y$ from both sides of the preceding equation, and recognizing that the inflation rate π is equal to the growth rate of the price level $\%\Delta P$, we can write:

$$\pi = \%\Delta P = \%\Delta M + \%\Delta V - \%\Delta Y$$

Since we assume velocity is constant, its growth rate is zero, and so the quantity theory of money is also a theory of inflation:

$$\pi = \%\Delta M - \%\Delta Y \tag{6}$$

Because the percentage change in a variable at an annual rate is the same as the growth rate of that variable, Equation 6 can be stated in words as follows: ***The quantity theory***

of inflation indicates that the inflation rate equals the growth rate of the money supply minus the growth rate of aggregate output. For example, if the aggregate output is growing at 3% per year and the growth rate of money is 5%, then inflation is 2%(= 5% − 3%). If the Bank of Canada increases the money growth rate to 10%, then the quantity theory of inflation given by Equation 6 indicates that the inflation rate will rise to 7%(= 10% − 3%).

APPLICATION Testing the Quantity Theory of Money

Now that we have fully outlined the quantity theory of money, let's put it to the test with actual data over the long and short runs.

The Quantity Theory of Money in the Long Run The quantity theory of money provides a long-run theory of inflation because it is based on the assumption that wages and prices are flexible. Panel (a) of Figure 20-1 plots 10-year averages of U.S. inflation rates against the 10-year average rate of U.S. money growth (M2) from 1870 through 2019. Because the growth rate of aggregate output Y over 10-year periods does not vary very much, Equation 6 indicates that the 10-year inflation rate should be the 10-year money growth rate minus a constant (the rate of aggregate output growth). Thus a strong positive relationship should exist between inflation and money growth rates—and this relationship is borne out in panel (a). Decades with higher growth rates of the U.S. money supply typically see higher average inflation rates.

Does the quantity theory also explain differing long-run inflation rates across countries? It certainly does. Panel (b) of Figure 20-1 plots the average inflation rate over the 10-year period from 2009 to 2019 against the 10-year money growth rate for several countries. Note that countries with high money growth rates, such as Russia and Turkey, tend to have higher inflation rates.

The Quantity Theory of Money in the Short Run Does the quantity theory of money provide a good explanation of short-run inflation fluctuations as well? Figure 20-2 provides evidence of the link between money growth and inflation in the short run by plotting the annual Canadian inflation rate from 1971 to 2020 against the annual money (M3) growth rate from two years before. (The money growth rate lags by two years to allow for the time it takes for changes in money growth to affect inflation.) The correlation between money growth and inflation on an annual basis is not strong at all. There were many years where money growth was high but inflation was low.

The conclusion from that data given in Figure 20-2 is that *the quantity theory of money is a good theory of inflation in the long run but not in the short run.* We could also say that *Milton Friedman's statement that "inflation is always and everywhere a monetary phenomenon" (mentioned in Chapter 1) is accurate in the long run but is not supported by the data for the short run.* This insight tells us that the classical assumption that wages and prices are completely flexible may not apply in the case of short-run fluctuations in inflation and aggregate output. For this reason, we relax this assumption in the following chapters of the book when we develop models of short-run inflation and output fluctuations. ◆

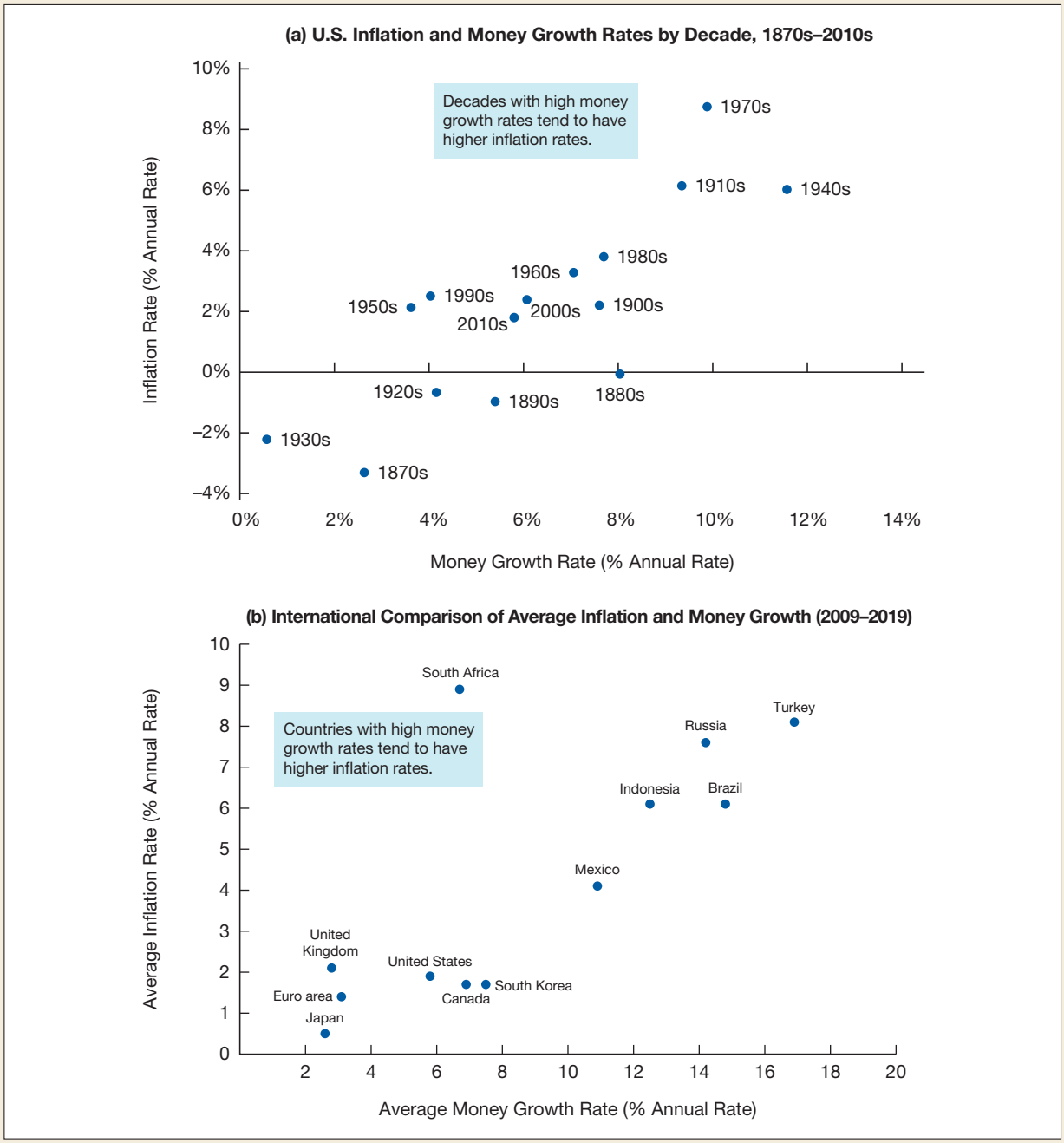


FIGURE 20-1 Relationship Between Inflation and Money Growth

In panel (a), decades with higher money growth rates (the 1910s, the 1940s, and the 1970s) typically have a higher average inflation rate. This relationship also holds in panel (b), where we examine the 10-year inflation and money growth rates from 2009 to 2019 for various countries.

Sources: Panel (a): Milton Friedman and Anna Schwartz, *Monetary Trends in the United States and the United Kingdom: Their Relation to Income, Prices, and Interest Rates, 1867–1975;* Federal Reserve Bank of St. Louis, FRED database: https://fred.stlouisfed.org/. Panel (b) Federal Reserve Bank of St. Louis, FRED database: https://fred.stlouisfed.org/.

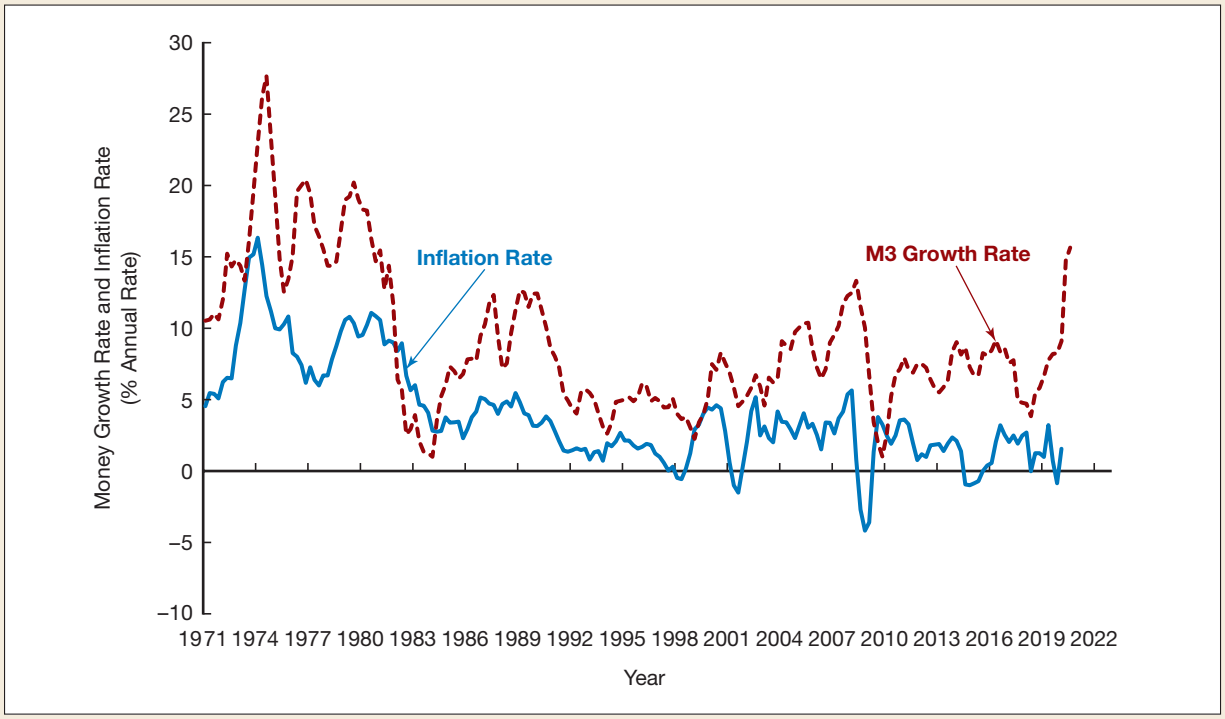


FIGURE 20-2 **Annual Canadian Inflation and Money Growth Rates, 1971–2020**

Plots of the annual Canadian inflation rate against the annual money (M3) growth rate from two years earlier (to allow for lag effects from money growth to inflation) do not support a short-run link between inflation and money growth. There were many years in which money growth was high yet inflation was low.

Source: Federal Reserve Bank of St. Louis, FRED (series CANGDPDEFQISMEI and MABMM301CAQ189S).

20.2 BUDGET DEFICITS AND INFLATION

LO 20.2 Identify the circumstances under which budget deficits can lead to inflationary monetary policy.

Budget deficits can be an important source of inflationary monetary policy. To see why this is the case, we need to look at the ways in which a government finances its budget deficits.

Government Budget Constraint

Because the government has to pay its bills just as we do, it has a budget constraint. We can pay for our spending in two ways: We can raise revenue (by working), or we can borrow. The government also enjoys these two options: It can raise revenue by levying

taxes, or it can go into debt by issuing government bonds. Unlike us, however, it has a third option: The government can create money and use it to pay for the goods and services it buys.

The methods used to finance government spending are described by an expression called the **government budget constraint**, which states the following: The government budget deficit *DEF*, which equals the excess of government spending *G* over tax revenue *T*, must equal the sum of the change in the monetary base ΔMB and the change in government bonds held by the public ΔB. Algebraically, this expression can be written as follows:

$$DEF = G - T = \Delta MB + \Delta B \tag{7}$$

To see what the government budget constraint means in practice, let's look at the case in which the only government purchase is a $100 million supercomputer. If the government convinces the electorate that such a computer is worth paying for, it will probably be able to raise the $100 million in taxes to pay for it, and the budget deficit will equal zero. According to the government budget constraint, no issue of money or bonds is needed to pay for the computer because the budget is balanced. If taxpayers think the supercomputer is too expensive and refuse to pay taxes to sponsor its purchase, the budget constraint indicates that the government must pay for the computer by selling $100 million of new bonds to the public or by, in effect, printing $100 million of currency. In either case, the budget constraint is satisfied. The $100 million deficit is balanced by the change in the stock of government bonds held by the public ($\Delta B = \$100$ million) or by the change in the monetary base ($\Delta MB = \$100$ million).

The government budget constraint thus reveals two important facts: ***If the government deficit is financed by an increase in bond holdings by the public, there is no effect on the monetary base and hence no effect on the money supply. But if the deficit is not financed by increased bond holdings by the public, both the monetary base and the money supply increase***.

There are several ways to understand why a deficit leads to an increase in the monetary base when the public's bond holdings do not increase. The simplest case is the case in which a government's treasury has the legal right to issue currency to finance its deficit. Financing the deficit is then very straightforward: The government just pays for the spending that is in excess of its tax revenues by issuing new currency. Because this increase in currency adds directly to the monetary base, the monetary base rises, and the money supply rises with it through the process of multiple deposit creation described in Chapter 15.

In Canada, however, and in many other countries, the government does not have the right to issue currency to pay for its bills. In this case, the government must finance its deficit by first issuing bonds to the public. If these bonds do not actually end up in the hands of the public, however, the only alternative is for the central bank to purchase them. To keep the newly issued government bonds from ending up in the hands of the public, the central bank must conduct an open market purchase, which, as we saw in Chapter 15, leads to an increase in the monetary base and in the money supply through the process of multiple deposit creation. This method of financing government spending is called **monetizing the debt** because, as indicated by the two-step process just described, government debt issued to finance government spending has been removed from the hands of the public and has been replaced

FYI Modern Monetary Theory

Modern Monetary Theory argues that the Green New Deal can be easily paid for by having the central bank buy government bonds to fund the resulting large budget deficits. As our analysis here suggests, Modern Monetary Theory is correct that the large budget deficit resulting from a large increase in government spending can be paid for by a central bank's purchases of government bonds. However, it ignores that this financing of a persistent budget deficit by a central bank's purchase of government bonds leads to an expansion in the money supply. The result of massive government spending financed by central bank purchases of government bonds can thus be very rapid growth of the money supply. As the quantity theory of money indicates, this very rapid growth of the money supply would then cause very high inflation.

Financing of the Green New Deal by the central bank's purchase of government bonds would thus not be "free" because it would likely result in very high inflation. Most mainstream economists, even those on the left, such as Paul Krugman, have rejected Modern Monetary Theory along the lines discussed above. Their criticisms are much in keeping with the economics adage "There is no such thing as a free lunch." A Green New Deal would need to be paid for by higher taxes at some point in the future if inflation is to be avoided and so should be evaluated on whether it is productive spending that will pay for itself or have sufficient benefits for the society that make it worth the cost.

by high-powered money. This method of financing is somewhat inaccurately referred to as **printing money** because high-powered money (the monetary base) is created in the process. The use of the word *printing* is misleading because no new currency is actually printed; instead, the monetary base increases when the central bank conducts open market purchases, just as it would increase if more currency were put into circulation.

We thus see that a budget deficit can lead to an increase in the money supply if it is financed by the creation of high-powered money. However, because the quantity theory of money explains inflation only in the long run, in order to produce inflation, the budget deficit must be *persistent*—that is, it must last for a substantial period of time. This leads us to the following conclusion: ***The financing of a persistent deficit by means of money creation will lead to sustained inflation.***

Progressives from the Democratic Party in the United States have recently proposed a "Green New Deal," a massive government spending program to cope with climate change and economic inequality. They have argued that a new theory referred to as *Modern Monetary Theory* (discussed in the FYI box "Modern Monetary Theory") makes it easy to pay for the Green New Deal.

Hyperinflation

The analysis here can be used to explain **hyperinflations**, periods of extremely high inflation of more than 50% per month. Many economies—both poor and developed—have experienced hyperinflation over the past century, but Canada has been spared such turmoil. One of the most extreme examples of hyperinflation throughout world history occurred in Zimbabwe in the 2000s, and it is discussed in the application that follows.

APPLICATION The Zimbabwean Hyperinflation

We now use our analysis of the quantity theory of money to explain the Zimbabwean hyperinflation that started in the early 2000s.

After the government expropriation of farms in 2000, which were redistributed to supporters of Robert Mugabe, the president of the country, Zimbabwean agricultural output plummeted and, along with it, tax revenue. As a result, the government's expenditures now massively exceeded revenues. The government could have obtained revenues to cover its expenditures by raising taxes, but given the depressed state of the economy, generating revenue in this way was both hard to do and would have been politically unpopular. Alternatively, the government could have tried to finance its expenditure by borrowing from the public, but given the public's distrust of the government, this was not an option. There was only one route left: the printing press. The government could pay for its expenditures by simply printing more currency (increasing the money supply) and using it to make payments to individuals and businesses. This is exactly what the Zimbabwean government did, and the money supply began to increase rapidly.

As predicted by the quantity theory, the surge in the money supply led to a rapidly rising price level. In February 2007, the Reserve Bank of Zimbabwe, the central bank, outlawed price increases on many commodities. Although this tactic has been tried before by governments in countries experiencing hyperinflations, it has never worked: Criminalizing inflation cannot stop inflation when the central bank keeps on printing money. In March 2007, the inflation rate hit a record of over 1500%. By 2008, Zimbabwe's official inflation rate was officially over 2 million percent (but unofficially over 10 million percent). In July 2008, the Zimbabwean central bank issued a new $100 billion bank note and soon after issued a $100 trillion dollar bill, the highest denomination dollar note on record. That's a lot of zeros, but don't be too impressed. Although holding one of these bills made you a trillionaire, such a bill could not even buy you a bottle of beer. Zimbabwean currency became worth less than toilet paper.

In 2009, the Zimbabwean government allowed the use of foreign currencies like the U.S. dollar for all transactions, but the damage had already been done. The hyperinflation wreaked havoc on the economy, and an extremely poor country became even poorer. ◆

20.3 KEYNESIAN THEORIES OF MONEY DEMAND

LO 20.3 Summarize the three motives underlying the liquidity preference theory of money demand.

In his famous 1936 book *The General Theory of Employment, Interest and Money*, John Maynard Keynes abandoned the quantity theory view that velocity is a constant and developed a theory of money demand that emphasized the importance of interest rates. In his theory of the demand for money, which he called the **liquidity preference theory**, Keynes presented three motives behind the demand for money: the transactions motive, the precautionary motive, and the speculative motive.

Transactions Motive

In the quantity theory approach, individuals are assumed to hold money because it is a medium of exchange that can be used to carry out everyday transactions. Keynes initially accepted the quantity theory view that the transactions component is proportional to income. Later, he and other economists recognized that new methods of payment, referred to as **payment technology**, could also affect the demand for money. For example, credit cards enable consumers to make even very small purchases without needing to hold money. Electronic payments that can be made from investors' brokerage accounts also reduce money demand. In Keynes's view, as payment technology advanced, the demand for money would be likely to decline relative to income.

Precautionary Motive

Keynes also recognized that people hold money as a cushion against unexpected opportunities. Suppose you have been thinking about buying a new Xbox gaming console and now see that it is on sale at 25% off. If you are holding money as a precaution for just such an occurrence, you can immediately buy it. Keynes argued that the precautionary money balances people would want to hold would also be proportional to income.

Speculative Motive

Keynes also believed that people choose to hold money as a store of wealth, which he called the *speculative motive*. Because the definition of money in Keynes's analysis includes currency (which earns no interest) and chequing account deposits (which typically earn little interest), he assumed that money earns no interest and hence its opportunity cost relative to holding other assets, such as bonds, is the nominal interest rate on bonds, i. As the interest rate i rises, the opportunity cost of money rises (it becomes more costly to hold money relative to bonds), and the quantity of money demanded falls.

Putting the Three Motives Together

In combining the three motives for holding money balances into a demand-for-money equation, Keynes was careful to distinguish between nominal quantities and real quantities. Money is valued in terms of what it can buy. If, for example, all prices in the economy doubled (the price level doubled), the same nominal quantity of money would be able to buy only half as many goods. Keynes thus reasoned that people want to hold a certain amount of **real money balances** (the quantity of money in real terms). By combining the three motives for holding money balances into a demand for real money balances, Keynes formulated what is called the liquidity preference function, which is written as follows:

$$\frac{M^d}{P} = \underset{-}{L}(\underset{-}{i}, \underset{+}{Y}) \tag{8}$$

Equation 8 states that the demand for real money balances is negatively related to the nominal interest rate and is positively related to real income.

Later Keynesian economists, such as Nobel Prize winner James Tobin, expanded the analysis and showed that interest rates play a more important role in money demand

than even Keynes supposed. These economists demonstrated that even the transactions and precautionary demands for money would be negatively related to the interest rate.[3]

An important implication of Keynesian theories of money demand is that velocity is not a constant but will fluctuate with changes in interest rates. To illustrate, we write the liquidity preference function as follows:

$$\frac{P}{M^d} = \frac{1}{L(i, Y)}$$

Multiplying both sides of this equation by Y and recognizing that we can replace M^d with M (because they must be equal in money market equilibrium), we can solve for velocity:

$$V = \frac{PY}{M} = \frac{Y}{L(i, Y)} \tag{9}$$

We know that the demand for money is negatively related to interest rates; when i goes up, $L(i, Y)$ declines, and therefore velocity rises. Because interest rates undergo substantial fluctuations, Keynesian theories of the demand for money indicate that velocity undergoes substantial fluctuations as well. Thus Keynesian theories cast doubt on the classical quantity theory view that nominal income is determined primarily by movements in the quantity of money.

20.4 PORTFOLIO THEORIES OF MONEY DEMAND

LO 20.4 Identify the factors underlying the portfolio choice theory of money demand.

Related to Keynes's analysis of the demand for money are so-called portfolio theories of money demand, in which people decide how much of an asset such as money they want to hold as part of their overall portfolio of assets.[4]

Theory of Portfolio Choice and Keynesian Liquidity Preference

In Chapter 5, we developed the theory of portfolio choice, which stated that the demand for an asset is positively related to wealth, the expected return on the asset relative to other assets, and relative liquidity, whereas it is negatively related to its risk relative to other assets. This theory of portfolio choice can justify the conclusion from the Keynesian liquidity preference function that the demand for real money balances is positively related to income and negatively related to the nominal interest rate.

Because income and wealth tend to move together, when income is higher, wealth is likely to be as well. Hence, higher income means greater wealth, and the theory of

[3]Three famous papers that elaborated on Keynes's approach to the demand for money are as follows: William J. Baumol, "The Transactions Demand for Cash: An Inventory Theoretic Approach," *Quarterly Journal of Economics* 66 (1952): 545–556; James Tobin, "The Interest Elasticity of the Transactions Demand for Cash," *Review of Economics and Statistics* 38 (1956): 241–247; and James Tobin, "Liquidity Preference as Behavior Towards Risk," *Review of Economic Studies* 25 (1958): 65–86.

[4]This is the approach taken by Milton Friedman in his famous paper "The Quantity Theory of Money: A Restatement," in *Studies in the Quantity Theory of Money*, ed. Milton Friedman (Chicago: University of Chicago Press, 1956), 3–21.

portfolio choice then indicates that the demand for the money assets will rise and the demand for real money balances will be higher.

As interest rates rise, the expected return on money does not change. However, the return on bonds, an alternative asset, goes up. Thus, although the expected *absolute* return on money did not change, money's expected return *relative* to bonds went down. In other words, as the theory of portfolio choice indicates, higher interest rates make money less desirable, and the demand for real money balances falls.

Other Factors That Affect the Demand for Money

The theory of portfolio choice indicates that other factors besides income and the nominal interest rate can affect the demand for money. We look at each of these in turn.

Wealth The theory of portfolio choice posits that as wealth increases, investors have more resources with which to purchase assets, increasing the demand for money. However, when income is held constant, greater wealth has only a small effect on the demand for money. In general, investors will hold only a small amount of money in their investment portfolio, preferring interest-bearing assets with similar risk and liquidity profiles, such as money market mutual funds, that are not included in measures of money such as M1. Currency and chequable deposits are sometimes said to be **dominated assets** because investors can hold other assets that pay higher returns and yet are perceived to be just as safe.

Risk It's hard to imagine an asset less risky than money. Currency will always be accepted, unless there's a revolution and the new government does not accept the old government's currency. And bank deposits are safe as long as deposit insurance exists. In the theory of portfolio choice, however, risk is always measured relative to another asset. Thus, if the stock market becomes more volatile, money can become less risky relative to stocks and demand for it will increase. In addition, although money is extremely safe on a nominal basis, its real return (the nominal return minus expected inflation) can become highly variable when inflation becomes very variable. Higher variability in the real return of money lowers the demand for money, as people shift into alternative assets known as **inflation hedges**, whose real returns are less affected than that of money when inflation varies. Popular inflation hedges include Real Return Bonds (discussed in Chapter 4), gold, and real estate.

Liquidity of Other Assets In recent years, financial innovation has led to the development of new liquid assets, such as money market mutual funds or home equity lines of credit that allow households to write cheques that are backed by their homes. As these alternative assets become more liquid, the relative liquidity of money falls, and so the demand for money falls as well.

Summary

Our analysis of the demand for money using Keynesian and portfolio theories indicates that seven factors affect the demand for money: interest rates, income, payment technology, wealth, riskiness of other assets, inflation risk, and liquidity of other assets. As a study aid, Summary Table 20-1 indicates the response of money demand to changes in each of these factors and gives a brief synopsis of the reasoning behind each response.

SUMMARY TABLE 20-1

Factors That Determine the Demand for Money

Variable	Change in Variable	Money Demand Response	Reason
Interest rates	↑	↓	Opportunity cost of money rises
Income	↑	↑	Higher value of transactions
Payment technology	↑	↓	Less need for money in transactions
Wealth	↑	↑	More resources to put into money
Riskiness of other assets	↑	↑	Money relatively less risky and so more desirable
Inflation risk	↑	↓	Money relatively more risky and so less desirable
Liquidity of other assets	↑	↓	Money relatively less liquid and so less desirable

Note: Only increases (↑) in the factors are shown; the effects of decreases in the variables on the exchange rate are the opposite of those indicated in the "Response" column.

20.5 EMPIRICAL EVIDENCE ON THE DEMAND FOR MONEY

LO 20.5 Assess and interpret the empirical evidence on the validity of the liquidity preference and portfolio theories of money demand.

Here we examine the empirical evidence on the two key issues that distinguish different theories of money demand and affect their conclusions about whether the quantity of money is the primary determinant of aggregate spending: Is the demand for money sensitive to changes in interest rates, and is the demand-for-money function stable over time?

Interest Rates and Money Demand

We have established that if interest rates do not affect the demand for money, velocity is more likely to be constant—or at least predictable—and so the quantity theory view that aggregate spending is determined by the quantity of money is more likely to be true. However, the more sensitive to interest rates the demand for money is, the more unpredictable velocity will be and the less clear the link between the money supply and aggregate spending becomes. Indeed, there exists an extreme case of ultrasensitivity of the demand for money to interest rates, called the **liquidity trap**, in which conventional monetary policy has no direct effect on aggregate spending because a change in the money supply has no effect on interest rates.[5]

[5]If the demand for money is ultrasensitive to interest rates, a tiny change in interest rates produces a very large change in the quantity of money demanded. Hence, in this case, the demand for money would be completely flat in the supply and demand diagrams of Chapter 5. Therefore, a change in the money supply that shifts the money supply curve to the right or left causes it to intersect the flat money demand curve at the same unchanged interest rate.

The evidence for the interest sensitivity of the demand for money is remarkably consistent. In situations in which nominal interest rates have not hit a floor of zero, the demand for money is sensitive to interest rates. However, when nominal interest rates fall to zero, they can go no lower. In this situation, a liquidity trap occurs because the demand for money is now completely flat. Indeed, this is exactly the situation that has occurred in Canada and the United States in recent years, which is why their central banks have had to resort to nonconventional monetary policy.

Stability of Money Demand

If the money demand function, like the one in Equation 8, is unstable and undergoes substantial, unpredictable shifts, as Keynes believed, then velocity is unpredictable, and the quantity of money may not be tightly linked to aggregate spending, as it is in the quantity theory. The stability of the money demand function is crucial to whether the Bank of Canada should target interest rates or the money supply. If the money demand function is unstable and so the money supply is not closely linked to aggregate spending, then the level of interest rates set by the Bank of Canada will provide more information about the stance of monetary policy than will the money supply.

Until the early 1970s, the evidence strongly supported the stability of the money demand function. However, after 1973, the rapid pace of financial innovation, which changed the items that could be used as money, led to substantial instability in estimated money demand functions. The instability of the money demand function calls into question whether our theories and empirical analyses are adequate. It also has important implications for the conduct of monetary policy, because it casts doubt on the usefulness of the money demand function as a tool for providing guidance to policymakers. In particular, because the money demand function has become unstable, velocity is now harder to predict. Monetary policymakers have found that the money supply does not provide reliable information on the future course of the economy, leading them to think of monetary policy in terms of the setting of interest rates. The instability of money demand has thus led to a downgrading of the focus on money supply in the conduct of monetary policy.

SUMMARY

1. The quantity theory of money as expressed by the equation of exchange, $M \times V = P \times Y$, indicates that nominal spending is determined solely by movements in the quantity of money. The quantity theory indicates that (1) changes in the quantity of money lead to proportional changes in the price level, because $P = (M \times \overline{V})/\overline{Y}$, and (2) the inflation rate is the growth rate of the money supply minus the growth rate of aggregate output—that is, $\pi = \%\Delta M - \%\Delta Y$. These implications of the quantity theory are borne out by the data in the long run but not in the short run.
2. The government budget constraint indicates that a deficit must be financed by either money creation or by the issuance of government bonds. That is, $DEF = G - T = \Delta MB + \Delta B$. Combining this fact with the quantity theory indicates that financing a persistent deficit by creating money will lead to sustained inflation. This analysis helps explain hyperinflations, in which inflation and money growth rise to extremely high levels because of massive budget deficits.
3. John Maynard Keynes suggested three motives for holding money: the transactions motive, the precautionary motive, and the speculative motive. His resulting liquidity preference theory views the transactions and precautionary components of money demand as proportional to income. However, the speculative

component of money demand is viewed as sensitive to interest rates as well as to expectations about the future movements of interest rates. This theory, then, implies that velocity is unstable and cannot be treated as a constant.

4. Portfolio theories of money demand indicate that the demand for money is determined not only by interest rates, income, and payment technology, as in the Keynesian analysis, but also by wealth, riskiness of other assets, inflation risk, and liquidity of other assets.
5. Two main conclusions can be reached from the research on the demand for money: The demand for money is sensitive to interest rates, as long as the interest rate is above zero. Since 1973, money demand has been found to be unstable, with the most likely source of the instability being the rapid pace of financial innovation. Because the money demand function is found to be both unstable and sensitive to interest rates, velocity cannot be viewed as constant and is not easily predicted. These conclusions have led to a downgrading of the focus on money supply and a greater emphasis on interest rates in the conduct of monetary policy.

MyLab Economics Make the grade with MyLab Economics: Select end-of-chapter questions can be found on MyLab Economics. You can practise them as often as you want, and most feature step-by-step guided instructions to help you find the right answer.

QUESTIONS

1. How would you expect velocity to typically behave over the course of the business cycle?
2. If velocity and aggregate output are reasonably constant (as the classical economists believed), what will happen to the price level when the money supply increases from $1 trillion to $4 trillion?
3. If credit cards were made illegal by government legislation, what would happen to velocity? Explain your answer.
4. "If nominal GDP rises, velocity must rise." Is this statement true, false, or uncertain? Explain your answer.
5. Why would a central bank be concerned about persistent, long-term budget deficits?
6. "Persistent budget deficits always lead to higher inflation." Is this statement true, false, or uncertain? Explain your answer.
7. Why might a central bank choose to monetize the debt, knowing that it could lead to higher inflation?
8. Consider two central banks: one with a history of maintaining price stability and low inflation and the other with a history of high inflation and poor inflation management. All else equal, if the same level of government budget deficit is monetized in both countries, how is inflation likely to behave in each country?
9. Some payment technologies require infrastructure (e.g., merchants need to have access to credit card swiping machines). In most developing countries historically this infrastructure has either been nonexistent or very costly. However, recently mobile payment systems have expanded rapidly in developing countries as they have become cheaper. Everything else being equal, would you expect the transaction component of the demand for money to be increasing or decreasing in a developing country relative to a rich country?
10. What three motives for holding money did Keynes consider in his liquidity preference theory of the demand for real money balances? On the basis of these motives, what variables did he think determined the demand for money?
11. In many countries, people hold money as a cushion against unexpected needs arising from a variety of potential scenarios (e.g., banking crises, natural

disasters, health problems, unemployment, etc.) that are not usually covered by insurance markets. Explain the effect of such behaviour on the precautionary component of the demand for money.

12. Why is Keynes's analysis of the speculative demand for money important to his view that velocity will undergo substantial fluctuations and thus cannot be treated as constant?

13. According to the portfolio theories of money demand, what are the four factors that determine money demand? What changes in these factors can increase the demand for money?

14. Explain how the following events will affect the demand for money according to the portfolio theories of money demand:

a. The economy experiences a business cycle contraction.

b. Brokerage fees decline, making bond transactions cheaper.

c. The stock market crashes. (*Hint:* Consider both the increase in stock price volatility following a market crash and the decrease in wealth of stockholders.)

15. Suppose a given country experienced low and stable inflation rates for quite some time, but then inflation picked up and over the past decade had been relatively high and quite unpredictable. Explain how this new inflationary environment would affect the demand for money according to portfolio theories of money demand. What would happen if the government decided to issue inflation-protected securities?

16. Consider the portfolio choice theory of money demand. How do you think the demand for money would be affected during a hyperinflation (i.e., monthly inflation rates in excess of 50%)?

17. Both the portfolio choice and Keynes's theories of the demand for money suggest that as the relative expected return on money falls, demand for it will fall. Why does the portfolio choice approach predict that money demand is affected by changes in interest rates? Why did Keynes think that money demand *is* affected by changes in interest rates?

18. Why does the Keynesian view of the demand for money suggest that velocity is unpredictable?

19. What evidence is used to assess the stability of the money demand function? What does the evidence suggest about the stability of money demand, and how has this conclusion affected monetary policymaking?

20. Suppose that a plot of the values of the money supply, M, and nominal GDP for a given country over 40 years shows that these two variables are very closely related. In particular, a plot of their ratio (nominal GDP/M) yields very stable and easy-to-predict values. On the basis of this evidence, would you recommend that the monetary authorities of this country conduct monetary policy by focusing mostly on the money supply rather than on setting interest rates? Explain.

APPLIED PROBLEMS

21. Suppose the money supply M has been growing at 10% per year, and nominal GDP, PY, has been growing at 20% per year. The data are as follows (in billions of dollars):

	2021	2022	2023
M	100	110	121
PY	1000	1200	1440

Calculate the velocity for each year. At what rate is the velocity growing?

22. Calculate what happens to nominal GDP if velocity remains constant at 5 and the money supply increases from \$200 billion to \$300 billion.

23. What happens to nominal GDP if the money supply grows by 20% but velocity declines by 30%?

24. If velocity and aggregate output remain constant at 5 and \$1000 billion, respectively, what happens to the price level if the money supply declines from \$400 billion to \$300 billion?

25. Suppose the liquidity preference function is given by

$$L(i, Y) = \frac{Y}{8} - 1000i$$

Use the money demand equation, along with the following table of values, to calculate the velocity for each period.

	Period 1	Period 2	Period 3	Period 4	Period 5	Period 6	Period 7
Y (in billions)	12 000	12 500	12 250	12 500	12 800	13 000	13 200
Interest rate	0.05	0.07	0.03	0.05	0.07	0.04	0.06

DATA ANALYSIS PROBLEMS

1. **Real-Time Data Analysis** The formula for computing the velocity of money is $Y/(M/P)$, where Y is real GDP, M is a money supply measure, and P is the price level. Go to the St. Louis FRED database and download quarterly Canadian data from 1981:Q1 to 2020:Q4 on the GDP deflator (series CANGDPDEFQISMEI) and M3 (series MABMM301CAQ189S). Also download nominal GDP from CANSIM (series V62295562).

   Compute the velocity of M3 money in 2020:Q3 and compare it to its level in 1981:Q1. Has it risen or fallen?

2. Use the data from the previous exercise and also download the three-month T-bill rate (CANSIM series V122531). Run a regression of real money balances on a constant, real GDP and the nominal interest rate.
   a. Comment on the fit of the regression.
   b. What is the estimated income elasticity of the demand for real money balances?
   c. What is the estimated interest rate elasticity? Are your estimates in (b) statistically significant?
   d. Is your evidence consistent with the Baumol-Tobin model or with the quantity theory of money?

3. **Real-Time Data Analysis** Go to the St. Louis Federal Reserve FRED database and find data on the M1 Money Stock (M1SL), M1 Money Velocity (M1V), and Real GDP (GDPC1). Convert the M1SL data series to "quarterly" using the *frequency* setting, and for all three series, use the "Percent Change from Year Ago" setting for *units*.
   a. Calculate the average percentage change in real GDP, the M1 money stock, and velocity since 2000:Q1.
   b. Based on your answer to part (a), calculate the average inflation rate since 2000 as predicted by the quantity theory of money.
   c. Next, find the data on the GDP deflator price index (GDPDEF), download the data using the "Percent Change from Year Ago" setting, and calculate the average inflation rate since 2000:Q1. Comment on the value relative to your answer in part (b).

21 The *IS* Curve

Learning Objectives

21.1 Explain the relationship between planned expenditure and aggregate demand.

21.2 List and describe the factors that determine the four components of aggregate demand (or planned expenditure).

21.3 Solve for the goods market equilibrium.

21.4 Describe why the *IS* curve slopes downward and why the economy heads to a goods market equilibrium.

21.5 List the factors that shift the *IS* curve, and describe how they shift the *IS* curve.

Preview

During the Great Depression of the 1930s, aggregate output fell precipitously, by 30%, with unemployment rising to 30%. Although the recession of 2007–2009 was not as severe, the contraction in economic activity led unemployment to rise to over 8.7%. With the coronavirus pandemic, the unemployment rate rose to a peak of 13.7% in May 2020. To understand why these contractions in economic activity occur, economists make use of the concept of *aggregate demand*, the total amount of output demanded in the economy. This concept was developed by John Maynard Keynes in his revolutionary book *The General Theory of Employment, Interest, and Money*, published in 1936, in which he argued that short-run changes in aggregate output, such as the decline in output that occurred during the Great Depression, are determined by changes in aggregate demand. The concept of aggregate demand is a central element in the *aggregate demand–aggregate supply (AD/AS) model*, the basic macroeconomic model used to explain short-run fluctuations in aggregate output.

In this chapter, we develop the first building block in understanding aggregate demand, the *IS curve*, which describes the relationship between real interest rates and aggregate output when the market for goods and services (more simply referred to as the *goods market*) is in equilibrium. We begin by deriving the *IS* curve and then go on to explain what factors cause the *IS* curve to shift. With our understanding of the *IS* curve, we can examine why fluctuations in economic activity occur and how the fiscal stimulus package of 2009 affected the economy. Then, in the next chapter, we make use of the *IS* curve to understand the role played by monetary policy in economic fluctuations.

21.1 PLANNED EXPENDITURE AND AGGREGATE DEMAND

LO 21.1 Explain the relationship between planned expenditure and aggregate demand.

We start our analysis by discussing the concept of **planned expenditure**, the total amount that households, businesses, the government, and foreigners want to spend on domestically produced goods and services. In contrast, *actual expenditure* is the amount that these entities actually do spend, which equals the total amount of output produced in the economy. Note that all of the analysis in this chapter refers to expenditure in real terms, that is, in terms of actual physical amounts of goods and services. Keynes viewed **aggregate demand**, the total amount of output demanded in the economy, as being

the same as planned expenditure. As we shall see shortly, planned expenditure—and hence aggregate demand—explains the level of aggregate output when the goods market is in equilibrium, that is, when aggregate demand for goods and services is equal to the actual amount of goods and services produced.

The total amount of aggregate demand (planned expenditure) is the sum of four types of spending:

1. **Consumption expenditure** (C), the total demand for consumer goods and services (e.g., hamburgers, iPhones, rock concerts, visits to the doctor)
2. **Planned investment spending** (I), the total planned spending by businesses on new physical capital (e.g., machines, computers, factories), plus planned spending on new homes
3. **Government purchases** (G), the spending by all levels of government on goods and services (e.g., aircraft carriers, salaries of government employees, red tape), not including transfer payments (which redistribute income from one person to another)
4. **Net exports** (NX), the net foreign spending on domestic goods and services, equal to exports minus imports

We represent the total aggregate demand (Y^{ad}) with the following equation:

$$Y^{ad} = C + I + G + NX \tag{1}$$

21.2 THE COMPONENTS OF AGGREGATE DEMAND

LO 21.2 List and describe the factors that determine the four components of aggregate demand (or planned expenditure).

To understand what determines aggregate demand (total planned expenditure) in the economy, let's look at each of its components in detail.

Consumption Expenditure

What determines how much you spend on consumer goods and services? Your income is likely the most important factor; if your income rises, you most likely will be willing to spend more. Keynes reasoned similarly that consumption expenditure is related to **disposable income** (denoted by Y_D), the total amount of income available for spending, equal to aggregate output Y minus taxes $T(Y - T)$.[1]

Consumption Function

Keynes called the relationship between disposable income Y_D and consumption expenditure C the **consumption function** and expressed it as follows:

$$C = \overline{C} + mpc \times Y_D \tag{2}$$

or, alternatively,

$$C = \overline{C} + mpc \times (Y - T) \tag{3}$$

[1]More precisely, taxes T refers to taxes minus net transfers (government payments to households and businesses that are, in effect, negative taxes). Examples of government transfers include Old Age Security pension payments and unemployment insurance payments.

FYI **Meaning of the Word *Investment***

Economists use the word *investment* somewhat differently than other people do. When noneconomists say that they are making an investment, they are normally referring to the purchase of common stocks or bonds, purchases that do not necessarily involve newly produced goods and services. But when economists speak of investment spending, they are referring to the purchases of *new* physical assets, such as new machines or new houses—purchases that add to aggregate demand.

The term $\overline{C}$ stands for **autonomous consumption expenditure**, the amount of consumption expenditure that is **exogenous** (independent of variables in the model, such as disposable income). Autonomous consumption is related to consumers' optimism about their future income and household wealth, both of which are positively related to consumer spending.

The term *mpc*, the **marginal propensity to consume**, reflects the change in consumption expenditure that results from an additional dollar of disposable income. Keynes assumed that *mpc* was a constant between the values of 0 and 1. If, for example, a $1.00 increase in disposable income leads to an increase in consumption expenditure of $0.60, then $mpc = 0.6$.

Planned Investment Spending

Investment spending is another key component of total expenditure. There are two types of investment spending: *fixed* and *inventory*. (Note that economists' use of the word *investment* differs from the everyday use of the term, as explained in the FYI box.)

Fixed Investment

Fixed investment is planned spending by firms on equipment (machines, computers, airplanes) and structures (factories, office buildings, shopping centres), plus planned spending on new residential housing.

Inventory Investment

Inventory investment is spending by firms on additional holdings of raw materials, parts, and finished goods, calculated as the change in holdings of these items in a given time period—say, a year.

Inventory investment is a much smaller component of investment than fixed investment. We discuss inventory investment in detail at this juncture because it plays an important role in the determination of aggregate output. To illustrate, consider the following scenarios:

1. Suppose that Ford Motor Company has 100 000 cars sitting in its factory lots on December 31, 2022, ready to be shipped to dealers. If each car has a wholesale price of $20 000, Ford has an inventory worth $2 billion. If by December 31, 2023, its inventory of cars has risen to 150 000, with a value of $3 billion, its inventory investment in 2023 is $1 billion, the *change* in the level of its inventory over the course of the year ($3 billion in 2023 minus $2 billion in 2022).
2. Instead suppose that by December 31, 2023, Ford's inventory of cars has dropped to 50 000, with a value of $1 billion. Its inventory investment in 2023 is now −$1 billion, the *change* in the level of its inventory over the course of the year ($1 billion in 2023 minus $2 billion in 2022).

3. Ford will have additional inventory investment if the level of raw materials and parts that it is holding to produce these cars increases over the course of the year. If on December 31, 2022, it holds $50 million of steel to be used to produce its cars, and on December 31, 2023, it holds $100 million of steel, it has an additional $50 million of inventory investment in 2023.

An important feature of inventory investment is that some inventory investment can be unplanned (in contrast, fixed investment is always planned). Suppose that the reason Ford finds itself with an additional $1 billion of cars on December 31, 2023, is because it sold $1 billion less of its cars than expected in 2023. This $1 billion of inventory investment in 2023 was unplanned. In this situation, Ford is producing more cars than it can sell, and it will cut production to avoid accumulating unsold goods. The act of adjusting production to eliminate unplanned inventory investment plays a key role in the determination of aggregate output, as we shall see.

Planned Investment Spending and Real Interest Rates Planned investment spending, a component of aggregate demand Y^{ad}, is equal to planned fixed investment plus the amount of inventory investment *planned* by firms. Keynes considered the level of the real interest rate for investments as a key determinant of planned investment spending.

To understand Keynes's reasoning, we need to recognize that businesses make investments in physical capital (machines and factories) as long as they expect to earn more from the physical capital than the interest cost of a loan to finance the investment. When the real interest rate for investments is high—say, at 10%—few investments in physical capital will earn more than the 10% interest cost of borrowing funds, and so planned investment spending will be low. When the real interest rate for investments is low—say, 1%—many investments in physical capital will earn more than the 1% interest cost of borrowing funds. Therefore, when the real interest rate for investments and hence the cost of borrowing are low, business firms are more likely to undertake investments in physical capital, and planned investment spending increases.

Even if a company has surplus funds and does not need to borrow to undertake an investment in physical capital, its planned investment spending still will be affected by the real interest rate for investments. Instead of investing in physical capital, the company could purchase a corporate bond. If the real interest rate on this security were high—say, 10%—the opportunity cost (forgone interest earnings) of an investment in physical capital would be high. Planned investment spending would then be low, because the firm probably would prefer to purchase the security and earn the high 10% return than to invest in physical capital. As the real interest rate for investments and the opportunity cost of investing fall—say, to 1%—planned investment spending will increase because investments in physical capital are likely to earn greater income for the firm than the measly 1% that would be earned by investing in the security.

Planned Investment and Business Expectations Keynes also believed that planned investment spending is heavily influenced by business expectations about the future. Businesses that are optimistic about future profit opportunities are willing to spend more, whereas pessimistic businesses cut back their spending. Thus Keynes posited a component of planned investment spending, which he called **autonomous investment**, $\bar{I}$, that is completely exogenous and so is unexplained by variables in his model, such as output or interest rates.

Keynes believed that changes in autonomous spending are dominated by these unstable exogenous fluctuations in planned investment spending, which are influenced

by emotional waves of optimism and pessimism—factors he labelled **"animal spirits."** His view was coloured by the collapse in investment spending during the Great Depression, which he saw as the primary reason for the economic contraction.

Investment Function By combining the two factors that Keynes theorized drive investment, we can derive an investment function that describes how planned investment spending is related to autonomous investment and the real interest rate for investments. We write this function as follows:

$$I = \bar{I} - dr_i \tag{4}$$

where d is a parameter reflecting the responsiveness of investment to the real interest rate for investments, which is denoted by r_i.

However, the real interest rate for investments reflects not only the real interest rate r on short-term, safe debt instruments, which is controlled by the central bank, but also **financial frictions**, denoted by $\bar{f}$, which are additions to the real cost of borrowing caused by barriers to the efficient functioning of financial markets. (We discussed the origins of these frictions—the asymmetric information problems of adverse selection and moral hazard—in detail in Chapter 8.) Financial frictions make it harder for lenders to ascertain the creditworthiness of a borrower. Lenders need to charge a higher interest rate to protect themselves against the possibility that the borrower may not pay back the loan, which leads to an increase in the *credit spread*, the difference between the interest rate on loans to businesses and the interest rate on completely safe assets that are sure to be paid back. Hence financial frictions add to the real interest rate for investments, and we can write:

$$r_i = r + \bar{f} \tag{5}$$

Substituting in Equation 4 the real cost of borrowing from Equation 5 yields:

$$I = \bar{I} - d(r + \bar{f}) \tag{6}$$

Equation 6 states that investment is positively related to business optimism as represented by autonomous investment and negatively related to the real interest rate and financial frictions.

Government Purchases and Taxes

Now we bring the government into the picture. The government affects aggregate demand in two ways: through its purchases and through taxes.

Government Purchases As we saw in the aggregate demand equation, Equation 1, government purchases add directly to aggregate demand. Here we assume that government purchases are exogenous, and we write government purchases as follows:

$$G = \overline{G} \tag{7}$$

Equation 7 states that government purchases are set at a fixed amount $\overline{G}$.

Taxes The government affects spending through taxes because, as discussed earlier, disposable income is equal to income minus taxes, $Y - T$, and disposable income affects consumption expenditure. Higher taxes T reduce disposable income for a given level of income and hence cause consumption expenditure to fall. The tax laws in a

country like Canada are very complicated, so to keep the model simple, we assume that government taxes are exogenous and are set at a fixed amount $\overline{T}$:[2]

$$T = \overline{T} \tag{8}$$

Net Exports

As with planned investment spending, we can think of net exports as being made up of two components: *autonomous net exports* and the part of net exports that is affected by changes in real interest rates.

Real Interest Rates and Net Exports Real interest rates influence the amount of net exports through the **exchange rate**. Recall that the exchange rate is the price of one currency, say, the dollar, in terms of another currency, say, the euro.[3] We examined a model that explains the link between the exchange rate and real interest rates in Chapter 18, but here we will only outline the intuition. When Canadian real interest rates rise, Canadian dollar assets earn higher returns relative to foreign assets. People then want to hold more Canadian dollars, so they bid up the value of a dollar and thereby increase its value relative to the values of other currencies. Thus a rise in Canadian real interest rates leads to a higher value of the Canadian dollar.

A rise in the value of the dollar makes Canadian exports more expensive in foreign currencies, so foreigners will buy less of these exports, thereby driving down net exports. A rise in the value of the dollar also makes foreign goods less expensive in terms of dollars, so Canadian imports will rise, also causing a decline in net exports. We therefore see that a rise in the real interest rate, which leads to an increase in the value of the dollar, in turn leads to a decline in net exports.

Autonomous Net Exports The amount of exports is also affected by the demand by foreigners for domestic goods, while the amount of imports is affected by the demand by domestic residents for foreign goods. For example, if the Chinese have a poor harvest and want to buy more Canadian wheat, Canadian exports will rise. If the Brazilian economy is booming, then Brazilians will have more money to spend on Canadian goods, and Canadian exports will rise. In contrast, if Canadian consumers discover how good Chilean wine is and want to buy more, then Canadian imports will rise. Thus we can think of net exports as being determined by real interest rates as well as by a component called **autonomous net exports**, $\overline{NX}$, which is the level of net exports that is treated as exogenous (outside the model).[4]

[2]For simplicity, we assume here that taxes are unrelated to income. However, because taxes increase with income, we can describe taxes more realistically with the following tax function:

$$T = \overline{T} + tY$$

Using this equation instead of Equation 9 in the derivation of Equation 12 later in the chapter would lead to *mpc* being replaced by $mpc(1 - t)$ in Equation 12.

[3]If a government pegs the exchange rate to another country's currency so that the rate is fixed in what is called a *fixed exchange rate regime* (see Chapter 19), then real interest rates do not directly affect net exports as in Equation 9, and $NX = \overline{NX}$. Taking out the response of net exports to the real interest rates does not change the basic analysis of the chapter but does lead to a slightly different Equation 12 later in the chapter.

[4]Foreign aggregate output is outside the model, and so its effect on net exports is exogenous and hence is a factor that affects autonomous net exports. Canadian domestic output, *Y*, could also affect net exports because greater domestic disposable income would increase spending on imports and thus would lower net exports. To build this factor into the *IS* curve, we could modify the net export function given in Equation 9 as follows:

$$NX = \overline{NX} - xr - iY$$

where *i* is the marginal propensity to spend on imports. This change would lead to a modification of Equation 12 later in the chapter, in which the *mpc* term would be replaced by $mpc - i$.

Net Export Function Putting these two components of net exports together, we can write a net export function:

$$NX = \overline{NX} - xr \tag{9}$$

where x is a parameter that indicates how net exports respond to the real interest rate. This equation tells us that net exports are positively related to autonomous net exports and are negatively related to the level of real interest rates.

21.3 GOODS MARKET EQUILIBRIUM

LO 21.3 Solve for the goods market equilibrium.

Keynes recognized that equilibrium will occur in the economy when the total quantity of output is equal to the total amount of aggregate demand (planned expenditure). That is,

$$Y = Y^{ad} \tag{10}$$

When this equilibrium condition is satisfied, planned spending for goods and services is equal to the amount that is produced. Producers are able to sell all of their output and have no reason to change their production levels, because there is no unplanned inventory investment. By examining the factors that affect each component of planned spending, we can understand why aggregate output goes to a certain level.

Solving for Goods Market Equilibrium

With our understanding of the factors that drive the components of aggregate demand, we can see how aggregate output is determined by using Equation 1, the aggregate demand equation, to rewrite the equilibrium condition given in Equation 10 as follows:

$$Y = C + I + G + NX \tag{11}$$

$$\begin{aligned}\text{aggregate output} = {} & \text{consumption expenditure} + \text{planned investment spending} \\ & + \text{government purchases} + \text{net exports}\end{aligned}$$

Now we can use our consumption, investment, and net export functions in Equations 3, 6, and 7, along with Equations 8 and 9, to determine aggregate output. Substituting all of these equations into the equilibrium condition given by Equation 11 yields the following:

$$Y = \overline{C} + mpc \times (Y - \overline{T}) + \overline{I} - d(r + \bar{f}) + \overline{G} + \overline{NX} - xr$$

Collecting terms, we can rewrite this equation as follows:

$$Y = \overline{C} + \overline{I} - d\bar{f} + \overline{G} + \overline{NX} + mpc \times Y - mpc \times \overline{T} - (d + x)r$$

Subtracting $mpc \times Y$ from both sides of the equation, we have

$$Y - mpc \times Y = Y(1 - mpc) = \overline{C} + \overline{I} - d\bar{f} + \overline{G} + \overline{NX} - mpc \times \overline{T} - (d + x)r$$

Then, dividing both sides of the equation by $1 - mpc$, we obtain an equation that gives us a means of determining aggregate output when the goods market is in equilibrium:[5]

$$Y = [\overline{C} + \overline{I} - d\overline{f} + \overline{G} + \overline{NX} - mpc \times \overline{T}] \times \frac{1}{1 - mpc} - \frac{d + x}{1 - mpc} \times r \quad (12)$$

Deriving the *IS* Curve

We refer to Equation 12 as the ***IS* curve**, and it shows the relationship between aggregate output and the real interest rate when the goods market is in equilibrium. Equation 12 is made up of two terms. Since *mpc* is between 0 and 1, $1/(1 - mpc)$ is positive, so the first term tells us that an increase in autonomous consumption, investment, government purchases, or net exports, or a decrease in taxes or financial frictions, leads to an increase in output at any given real interest rate. In other words, the first term tells us about shifts in the *IS* curve. The second term tells us that an increase in real interest rates results in a decrease in output, which can be shown as a movement along the *IS* curve.

21.4 UNDERSTANDING THE *IS* CURVE

LO 21.4 Describe why the *IS* curve slopes downward and why the economy heads to a goods market equilibrium.

To gain a deeper understanding of the *IS* curve, we will proceed in several steps. In this section, we begin by looking at the intuition behind the *IS* curve, and then we discuss a numerical example. Then, in the following section, we outline the factors that shift the *IS* curve.

What the *IS* Curve Tells Us: Intuition

The *IS* curve traces out the points at which the goods market is in equilibrium. For each given level of the real interest rate, the *IS* curve tells us the level of aggregate output that is necessary for the goods market to be in equilibrium. As the real interest rate rises, planned investment spending and net exports fall, which in turn lowers aggregate demand; aggregate output must be lower if it is to equal aggregate demand and satisfy goods market equilibrium. Hence the *IS* curve is downward-sloping.

What the *IS* Curve Tells Us: Numerical Example

We can analyze the *IS* curve with the following numerical example, which gives specific values for the exogenous variables and the parameters in Equation 12.

$$\begin{aligned} \overline{C} &= \$1.4 \text{ trillion} \\ \overline{I} &= \$1.2 \text{ trillion} \\ \overline{G} &= \$3.0 \text{ trillion} \\ \overline{T} &= \$3.0 \text{ trillion} \\ \overline{NX} &= \$1.3 \text{ trillion} \\ \overline{f} &= 1 \\ mpc &= 0.6 \\ d &= 0.3 \\ x &= 0.1 \end{aligned}$$

[5]Note that the term $1/(1 - mpc)$ that multiplies $\overline{G}$ is known as the *expenditure multiplier*, and the term $-\text{mpc}/(1 - mpc)$ that multiplies $\overline{T}$ is called the *tax multiplier*. The tax multiplier is smaller in absolute value than the expenditure multiplier because $mpc < 1$.

Mini-lecture

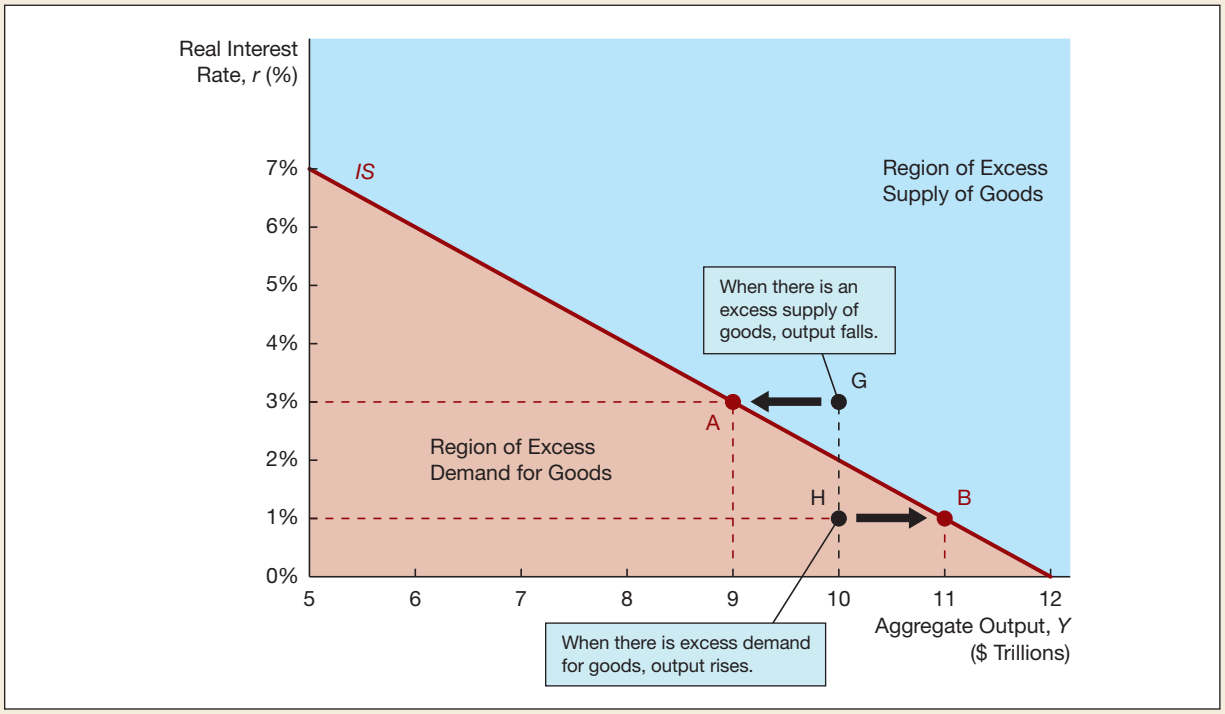


FIGURE 21-1 The *IS* Curve

The downward-sloping *IS* curve represents points at which the goods market is in equilibrium—for example, points A and B. Notice that output changes as necessary to return the market to equilibrium. For example, at point G in the blue-shaded area, an excess supply of goods exists, and firms will cut production, decreasing aggregate output to the equilibrium level at point A. At point H in the orange-shaded area, an excess demand for goods exists, so firms will increase production, and aggregate output will increase toward the equilibrium level at point B.

Using these values, we can rewrite Equation 12 as follows:

$$Y = [1.4 + 1.2 - 0.3 + 3.0 + 1.3 - 0.6 \times 3.0] \times \frac{1}{1 - 0.6} - \frac{0.3 + 0.1}{1 - 0.6} \times r$$

Plugging these values into Equation 12 yields the equation of the *IS* curve shown in Figure 21-1:

$$Y = \frac{4.8}{0.4} - \frac{0.4}{0.4} \times r = 12 - r \qquad (13)$$

At a real interest rate of $r = 3\%$, the equilibrium output Y is equal to \$12 trillion − \$3 trillion = \$9 trillion. We plot this combination of the real interest rate and equilibrium output as point A in Figure 21-1. At a real interest rate of $r = 1\%$, the equilibrium output Y is equal to \$12 trillion − \$1 trillion = \$11 trillion, which we plot as point B. The line connecting these points is the *IS* curve, and, as you can see, it is downward-sloping.

Why the Economy Heads Toward Equilibrium

The concept of equilibrium is useful only if there is a tendency for the economy to settle there. Let's first consider what happens if the economy is located to the right of the *IS* curve (the blue-shaded area), where an excess supply of goods exists. In Figure 21-1 at point G, actual output is above aggregate demand, and firms are saddled with unsold inventory. To keep from accumulating unsold goods, firms will continue cutting production. As long as production is above the equilibrium level, output will exceed aggregate demand and firms will continue cutting production, sending aggregate output toward the equilibrium level, as indicated by the leftward arrow from point G to point A. Only when the economy moves to point A on the *IS* curve will there be no further tendency for output to change.

What happens if aggregate output is below the equilibrium level of output (the orange-shaded area to the left of the *IS* curve), where an excess demand for goods exists? At point H in Figure 21-1, actual output is below aggregate demand, so firms will want to increase production because inventories are declining more than they desire, and aggregate output will increase, as shown by the rightward arrow. When the economy has moved to point B on the *IS* curve, there will again be no further tendency for output to change.

21.5 FACTORS THAT SHIFT THE *IS* CURVE

LO 21.5 List the factors that shift the *IS* curve, and describe how they shift the *IS* curve.

You have now learned that the *IS* curve describes equilibrium points in the goods market—the combinations of the real interest rate and equilibrium output. The *IS* curve shifts whenever change occurs in autonomous factors (factors independent of aggregate output and the real interest rate). Note that a change in the real interest rate that affects equilibrium aggregate output causes only a *movement along* the *IS* curve. A *shift* in the *IS* curve, by contrast, occurs when equilibrium output changes *at each given real interest rate*.

In Equation 12, we identified six autonomous factors that can shift aggregate demand and hence affect the level of equilibrium output. Although Equation 12 directly tells us how these factors shift the *IS* curve, we will develop some intuition as to how each autonomous factor does so.

Changes in Government Purchases

Let's look at what happens if government purchases rise from \$3 trillion to \$4 trillion in Figure 21-2. IS_1 represents the same *IS* curve that we developed in Figure 21-1. We determine the equation for IS_2 by substituting the \$4 trillion value into Equation 12:

$$Y = [1.4 + 1.2 - 0.3 + 4.0 + 1.3 - 0.6 \times 3.0] \times \frac{1}{1 - 0.6} - \frac{0.3 + 0.1}{1 - 0.6} \times r$$
$$= \frac{5.8}{0.4} - r = 14.5 - r$$

On the basis of these results, at a real interest rate of $r = 3\%$, equilibrium output Y is equal to \$14.5 trillion − \$3 trillion = \$11.5 trillion, which we mark as point C on Figure 21-2. At a real interest rate of $r = 1\%$, equilibrium output has increased to $Y =$ \$14.5 trillion − \$1 trillion = \$13.5 trillion, which we mark as point D. The increase in government purchases therefore shifts the *IS* curve to the right from IS_1 to IS_2.

Mini-lecture

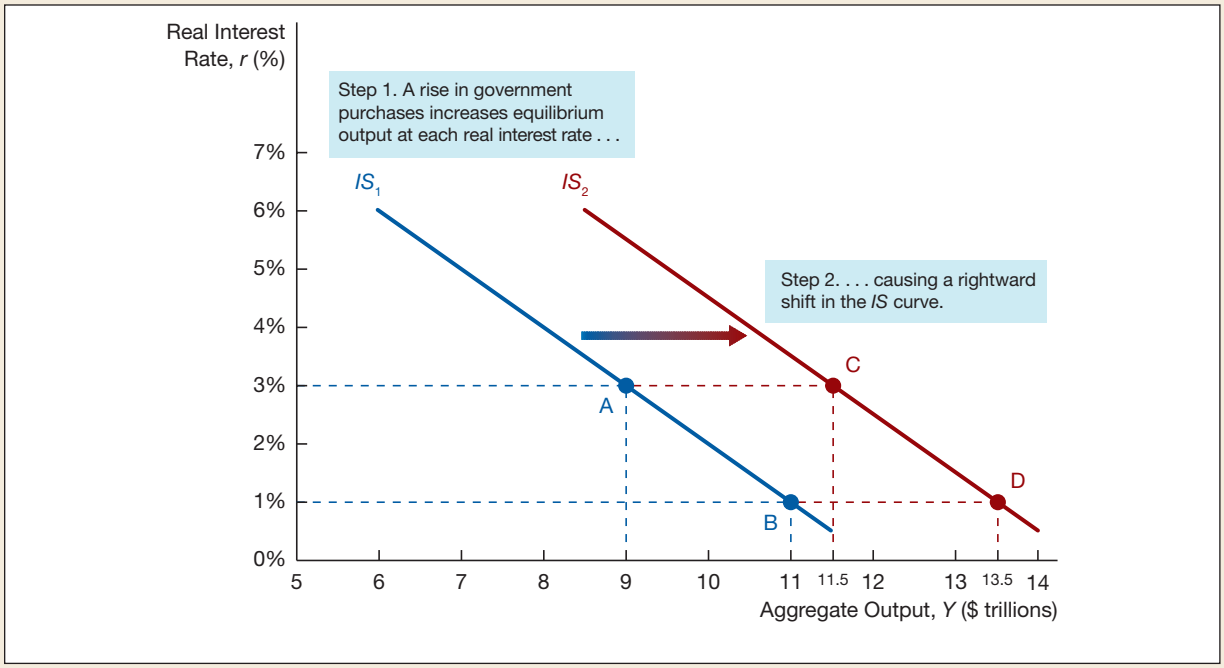


FIGURE 21-2 Shift in the *IS* Curve from an Increase in Government Purchases

IS_1 represents the *IS* curve we derived in Figure 21-1. IS_2 reflects a $1.0 trillion increase in government purchases. The increase in government purchases causes aggregate output to rise, shifting the *IS* curve to the right by $2.5 trillion at each real interest rate, from IS_1 to IS_2.

Intuitively, we can see why an increase in government purchases leads to a rightward shift of the *IS* curve by recognizing that an increase in government purchases causes aggregate demand to increase at any given real interest rate. Since aggregate output equals aggregate demand when the goods market is in equilibrium, ***an increase in government purchases that causes aggregate demand to rise also causes equilibrium output to rise, thereby shifting the* IS *curve to the right. Conversely, a decrease in government purchases causes aggregate demand to fall at any given real interest rate and leads to a leftward shift of the* IS *curve.***

APPLICATION The Vietnam War Buildup, 1964–1969

The United States' involvement in Vietnam began to escalate in the early 1960s. After 1964, the United States was fighting a full-scale war. Beginning in 1965, the resulting increases in military expenditure raised government purchases. When government purchases are rising rapidly, central banks will usually raise real interest rates to keep the economy from overheating. The Vietnam War period, however, was unusual in that the Federal Reserve decided to keep real interest rates constant. Hence, this period provides an excellent example of how policymakers might make use of *IS* curve analysis to inform policy.

Mini-lecture

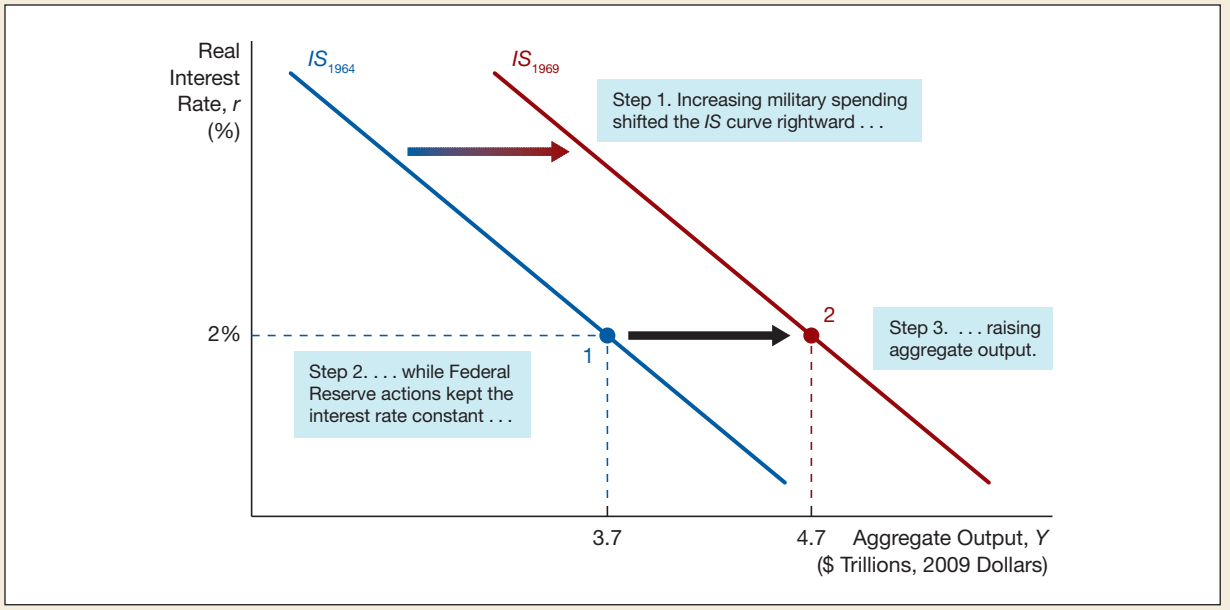

Real Interest Rate, r (%)
IS1964
IS1969
Step 1. Increasing military spending shifted the IS curve rightward . . .
2%
1
2
Step 2. . . . while Federal Reserve actions kept the interest rate constant . . .
Step 3. . . . raising aggregate output.
3.7
4.7
Aggregate Output, Y ($ Trillions, 2009 Dollars)

FIGURE 21-3 Vietnam War Buildup

Increases in military spending beginning in 1965 caused the *IS* curve to shift from IS_{1964} to IS_{1969}. Because the Federal Reserve decided to keep real interest rates constant at 2% during this period, equilibrium output rose from \$3.7 trillion (in 2009 dollars) in 1964 to \$4.7 trillion in 1969, setting the stage for an increase in inflation.

The rise in government purchases shifted the *IS* curve to the right, from IS_{1964} to IS_{1969} in Figure 21-3. Because the Federal Reserve decided to keep real interest rates constant at 2% during this period, equilibrium output rose from \$3.7 trillion (in 2009 dollars) in 1964 to \$4.7 trillion by 1969, with the unemployment rate falling steadily from 5% in 1964 to 3.4% in 1969. However, all was not well for the economy: The combination of an increase in government purchases and a constant real interest rate led to an overheating of the economy that eventually resulted in high inflation. (We will discuss the link between an overheating economy and inflation in the coming chapters.) ◆

Changes in Taxes

Now let's look at Figure 21-4 to see what happens if the government raises taxes from \$3 trillion to \$4 trillion. IS_1 represents the same *IS* curve that we developed in Figure 21-1. We determine the equation for IS_2 by substituting the \$4 trillion value into Equation 12:

$$Y = [1.4 + 1.2 - 0.3 + 3.0 + 1.3 - 0.6 \times 4.0] \times \frac{1}{1 - 0.6} - \frac{0.3 + 0.1}{1 - 0.6} \times r$$

$$= \frac{4.2}{0.4} - r = 10.5 - r$$

Mini-lecture

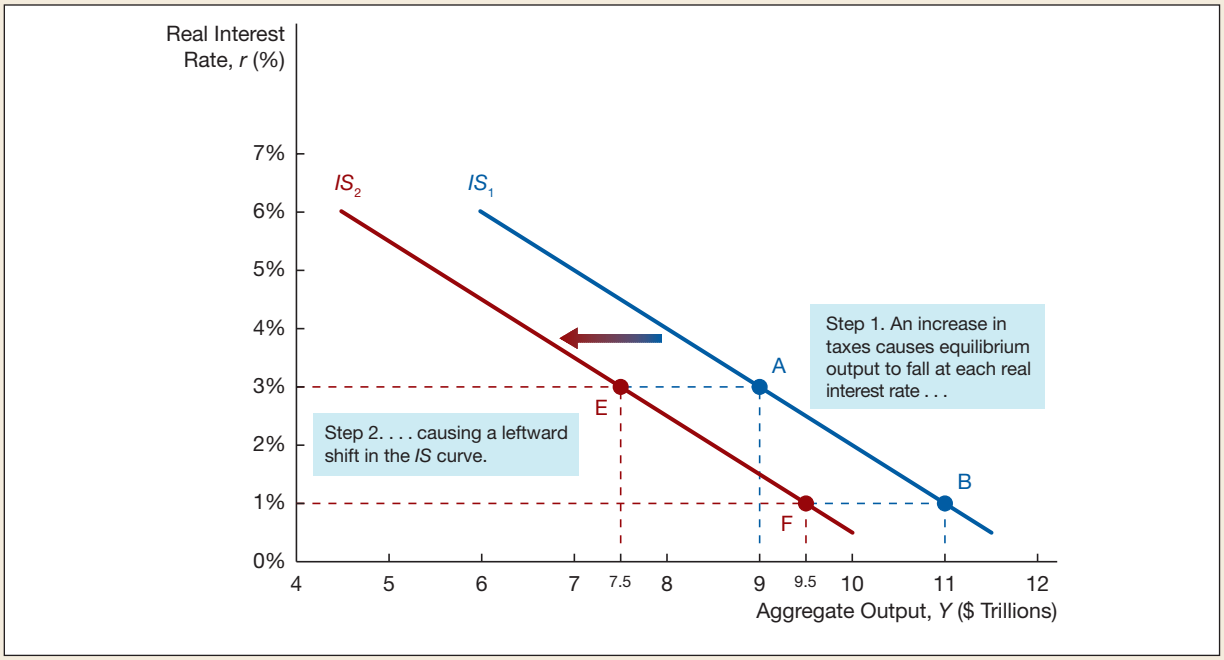


FIGURE 21-4 Shift in the *IS* Curve from an Increase in Taxes

IS_1 represents the *IS* curve we derived in Figure 21-1. IS_2 reflects a \$1.0 trillion increase in government tax revenues. The increase in taxes decreases aggregate output levels by \$1.5 trillion, shifting the *IS* curve to the left, from IS_1 to IS_2.

At a real interest rate of $r = 3\%$, equilibrium output $Y = \$10.5$ trillion $- \$3$ trillion $= \$7.5$ trillion, which we mark as point E in Figure 21-4. At this real interest rate, equilibrium output has decreased from point A to point E, as shown by the leftward arrow. Similarly, at a real interest rate of $r = 1\%$, equilibrium output has decreased to $Y = \$10.5$ trillion $- \$1$ trillion $= \$9.5$ trillion, causing a leftward shift from point B to point F. The *IS* curve shifts to the left, from IS_1 to IS_2, as a result of the increase in taxes.

We have the following result: ***At any given real interest rate, a rise in taxes causes aggregate demand and hence equilibrium output to fall, thereby shifting the* IS *curve to the left. Conversely, a cut in taxes at any given real interest rate increases disposable income and causes aggregate demand and equilibrium output to rise, shifting the* IS *curve to the right.***

Policymakers use both tax and government purchase policies to stimulate the economy when it enters a recession, as illustrated in the following application.

APPLICATION The U.S. Fiscal Stimulus Package of 2009

In the fall of 2008, the U.S. economy was in crisis. By the time the new Obama administration took office, the unemployment rate had risen from 4.7% just before the recession began in December 2007 to 7.6% in January 2009. To stimulate the economy, the

Obama administration proposed a fiscal stimulus package that, when passed by Congress, included $288 billion in tax cuts for households and businesses and $499 billion in increased federal spending, including transfer payments. What does our *IS* curve analysis suggest should have happened to the economy?

As the analyses in Figure 21-2 and Figure 21-4 indicate, these tax cuts and spending increases should have increased aggregate demand, thereby raising the equilibrium level of aggregate output at any given real interest rate and so shifting the *IS* curve to the right. Unfortunately, things didn't work out quite as the Obama administration had planned. Most of the government purchases did not kick in until after 2010, while the declines in autonomous consumption and investment were much larger than anticipated. The fiscal stimulus was more than offset by the weak consumption and investment caused by an increase in financial frictions and worries about the economy. As a result, aggregate demand ended up contracting rather than rising, and the *IS* curve did not shift to the right as hoped. Despite the good intentions of the fiscal stimulus package, the unemployment rate ended up rising to 10% in 2009. Without the fiscal stimulus package, however, the *IS* curve likely would have shifted even further to the left, resulting in even more unemployment. ◆

Changes in Autonomous Spending

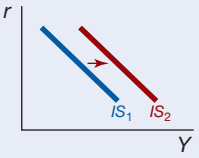
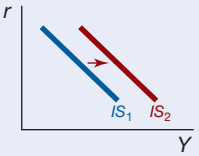
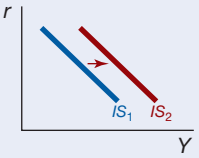
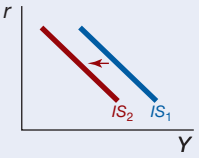
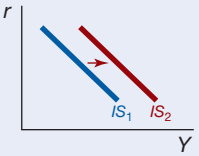
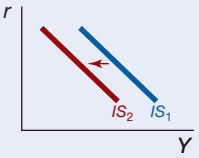
As you can see from Equation 12, autonomous consumption, investment, and net exports—$\overline{C}$, $\overline{I}$, and $\overline{NX}$, respectively—all are multiplied by the term $1/(1 - mpc)$ in the same way the $\overline{G}$ term is. Thus an increase in any of these variables has the same impact on the *IS* curve as an increase in government purchases. For this reason, we can lump these variables together as **autonomous spending**, exogenous spending that is unrelated to variables in the model such as output or real interest rates. We look intuitively at how changes in each of these variables affect the *IS* curve in turn.

Autonomous Consumption Suppose consumers find that their wealth has increased courtesy of a stock market boom or that they have become increasingly optimistic about their future income prospects because a positive productivity shock to the economy has occurred. Both of these events are autonomous; that is, they are not affected by the level of the real interest rate. ***The resulting rise in autonomous consumption would raise aggregate demand and equilibrium output at any given real interest rate, shifting the IS curve to the right. Conversely, a decline in autonomous consumption expenditure would cause aggregate demand and equilibrium output to fall, shifting the IS curve to the left.***

Autonomous Investment Spending Earlier in the chapter, we learned that changes in the real interest rate affect planned investment spending and hence the equilibrium level of output. This change in investment spending merely causes a movement along the *IS* curve, not a shift. An autonomous rise in planned investment spending unrelated to the real interest rate—say, because companies become more confident about investment profitability after the stock market rises—increases aggregate demand. ***An increase in autonomous investment spending therefore increases equilibrium output at any given real interest rate, shifting the IS curve to the right. In contrast, a decrease in autonomous investment spending causes aggregate demand and equilibrium output to fall, shifting the IS curve to the left.***

SUMMARY TABLE 21-1

Shifts in the *IS* Curve from Autonomous Changes in $\overline{C}$, $\overline{I}$, $\overline{G}$, $\overline{T}$, $\overline{NX}$, and $\overline{f}$

Variable	Change in Variable	Shift in *IS* Curve	Reason
Autonomous consumption expenditure, $\overline{C}$	↑		$C \uparrow Y \uparrow$
Autonomous investment, $\overline{I}$	↑		$I \uparrow Y \uparrow$
Government spending, $\overline{G}$	↑		$G \uparrow Y \uparrow$
Taxes, $\overline{T}$	↑		$T \uparrow \Rightarrow C \downarrow Y \downarrow$
Autonomous net exports, $\overline{NX}$	↑		$\overline{NX} \uparrow Y \uparrow$
Financial frictions, $\overline{f}$	↑		$I \downarrow Y \downarrow$

Note: Only increases (↑) in the variables are shown; the effects of decreases in the variables on aggregate output would be the opposite of those indicated in the last two columns.

Autonomous Net Exports An autonomous rise in net exports unrelated to the real interest rate—say, because Canadian-made handbags become more chic than French-made handbags, or because foreign countries have a boom and thus buy more Canadian goods—causes aggregate demand to rise. ***An autonomous increase in net exports thus leads to an increase in equilibrium output at any given real interest rate and shifts the* IS *curve to the right. Conversely, an autonomous fall in net exports causes aggregate demand and equilibrium output to decline, shifting the* IS *curve to the left.***

Changes in Financial Frictions

An increase in financial frictions, as occurred during the financial crisis of 2007–2009, raises the real interest rate for investments and hence causes investment spending and aggregate demand to fall. ***An increase in financial frictions leads to a decline in equilibrium output at any given real interest rate and shifts the* IS *curve to the left. Conversely, a decrease in financial frictions causes aggregate demand and equilibrium output to rise, shifting the* IS *curve to the right.***

Summary of Factors That Shift the *IS* Curve

As a study aid, Summary Table 21-1 shows how each factor shifts the *IS* curve and the reason the shift occurs. Now that we have a full understanding of the *IS* curve, we can use this building block to examine the relationship between monetary policy and the aggregate demand curve in the following chapter.

SUMMARY

1. Planned expenditure, the total amount of goods demanded in the economy, is the same as aggregate demand, which is the sum of four types of spending: consumption expenditure, planned investment spending, government purchases, and net exports. We represent the total aggregate demand (Y^{ad}) with Equation 1: $Y^{ad} = C + I + G + NX$.
2. Consumption expenditure is described by the consumption function, which indicates that consumption expenditure will rise as disposable income increases. Planned expenditure and hence aggregate demand are negatively related to the real interest rate because a rise in the real interest rate reduces both planned investment spending and net exports. An increase in financial frictions raises the real interest rate for investments and hence lowers planned investment spending and aggregate demand. The government also affects planned expenditure via spending, which directly changes aggregate demand, or via taxes, which indirectly affect aggregate demand by influencing disposable income and hence consumption expenditure.
3. The level of aggregate output when the goods market is in equilibrium is determined by the condition that aggregate output equals aggregate demand.
4. The *IS* curve traces out the combinations of the real interest rate and aggregate output at which the goods market is in equilibrium. The *IS* curve slopes downward because higher real interest rates lower planned investment spending and net exports and so lower equilibrium output.
5. The *IS* curve shifts to the right when there is a rise in autonomous consumption, a rise in autonomous investment, a rise in government purchases, a rise in autonomous net exports, a fall in taxes, or a decline in financial frictions. Movements of these six factors in the opposite direction will shift the *IS* curve to the left.

MyLab Economics Make the grade with MyLab Economics: Select end-of-chapter questions can be found on MyLab Economics. You can practise them as often as you want, and most feature step-by-step guided instructions to help you find the right answer.

QUESTIONS

1. "When the stock market rises, investment spending is increasing." Is this statement true, false, or uncertain? Explain your answer.
2. Why is inventory investment counted as part of aggregate spending if it isn't actually sold to the final end user?
3. "Since inventories can be costly to hold, firms' planned inventory investment should be zero, and firms should acquire inventory only through unplanned inventory accumulation." Is this statement true, false, or uncertain? Explain your answer.
4. During and in the aftermath of the financial crisis of 2007–2009, planned investment fell substantially despite significant decreases in the real interest rate. What factors related to the planned investment function could explain this?
5. If households and firms believe the economy will be in a recession in the future, will this necessarily cause a recession or have any impact on output at all?
6. Why do increases in the real interest rate lead to decreases in net exports and vice versa?
7. How would a decline in house prices or decrease in stock market prices affect the *IS* curve?
8. If firms suddenly become more optimistic about the profitability of investment and planned investment spending rises by $100 billion, while consumers become more pessimistic and autonomous consumer spending falls by $100 billion, what happens to aggregate output?
9. If an increase in autonomous consumer expenditure is matched by an equal increase in taxes, will aggregate output rise or fall?
10. If a change in the real interest rate has no effect on planned investment spending or net exports, what does this imply about the slope of the *IS* curve?
11. Inventories typically increase starting at the beginning of recessions and begin to decline near the end of recessions. What does this say about the relationship between planned spending and aggregate output over the business cycle?
12. Why do companies cut production when they find that their unplanned inventory investment is greater than zero? If they didn't cut production, what effect would this have on their profits? Why?
13. "Firms will increase production when planned investment is less than (actual) total investment." Is this statement true, false, or uncertain? Explain your answer.
14. In each of the following cases, determine whether the *IS* curve shifts to the right or left, does not shift, or is indeterminate in the direction of shift.
    a. The real interest rate rises.
    b. The marginal propensity to consume declines.
    c. Financial frictions increase.
    d. Autonomous consumption decreases.
    e. Both taxes and government spending decrease by the same amount.
    f. The sensitivity of net exports to changes in the real interest rate decreases.
    g. The government provides tax incentives for research and development programs for firms.
15. "Financial frictions are not a problem for the economy, since they do not affect the safe policy rate which is controlled by the central bank." Is this statement true, false, or uncertain? Explain your answer.
16. When the Bank of Canada reduces its policy interest rate, how, if at all, is the *IS* curve affected? Briefly explain.
17. Suppose you read that prospects for stronger future economic growth have led the Canadian dollar to strengthen and stock prices to increase.
    a. What effect does the strengthened dollar have on the *IS* curve?
    b. What effect does the increase in stock prices have on the *IS* curve?
    c. What is the combined effect of these two events on the *IS* curve?

APPLIED PROBLEMS

18. Calculate the value of the consumption function at each level of income in the following table if autonomous consumption = 300, taxes = 200, and *mpc* = 0.9.

Income Y	Disposable Income Y^D	Consumption C
0		
100		
200		
300		
400		
500		
600		

19. Assume that autonomous consumption is $1625 billion and disposable income is $11 500 billion. Calculate consumption expenditure if an increase of $1000 in disposable income leads to an increase of $750 in consumption expenditure.

20. Suppose that Dell Corporation has 20 000 computers in its warehouses on December 31, 2022, ready to be shipped to merchants (each computer is valued at $500). By December 31, 2023, Dell Corporation has 25 000 computers ready to be shipped, each valued at $450.

a. Calculate Dell's inventory on December 31, 2022.

b. Calculate Dell's inventory investment in 2023.

c. What happens to inventory spending during the early stages of an economic recession?

21. If the consumption function is $C = 100 + 0.75Y_D$, $I = 200$, government spending is 200, and net exports are zero, what will be the equilibrium level of output? What will happen to aggregate output if government spending rises by 100?

22. If the marginal propensity to consume is 0.75, by how much would government spending have to rise to increase output by $1000 billion? By how much would taxes need to decrease to increase output by $1000 billion?

23. Assuming both taxes and government spending increase by the same amount, derive an expression for the effect on equilibrium output.

24. Consider an economy described by the following data:

$\bar{C}$ = $3.25 trillion
$\bar{I}$ = $1.3 trillion
$\bar{G}$ = $3.5 trillion
$\bar{T}$ = $3.0 trillion
$\overline{NX}$ = −$1.0 trillion
$\bar{f}$ = 1
mpc = 0.75
d = 0.3
x = 0.1

a. Derive simplified expressions for the consumption function, the investment function, and the net export function.

b. Derive an expression for the *IS* curve.

c. If the real interest rate is $r = 2$, what is equilibrium output? If $r = 5$, what is equilibrium output?

d. Draw a graph of the *IS* curve showing the answers from part (c) above.

e. If government purchases increase to $4.2 trillion, what will happen to equilibrium output at $r = 2$? What will happen to equilibrium output at $r = 5$? Show the effect of the increase in government purchases in your graph from part (d).

25. Consider an economy described by the following data:

$\bar{C}$ = $4 trillion
$\bar{I}$ = $1.5 trillion
$\bar{G}$ = $3.0 trillion
$\bar{T}$ = $3.0 trillion
$\overline{NX}$ = $1.0 trillion
$\bar{f}$ = 0
mpc = 0.8
d = 0.35
x = 0.15

a. Derive an expression for the *IS* curve.

b. Assume that the Bank of Canada controls the interest rate and sets the interest rate at $r = 4$. What is the equilibrium level of output?

c. Suppose that a financial crisis begins and $\bar{f}$ increases to $\bar{f} = 3$. What will happen to equilibrium output?

If the Bank of Canada can set the interest rate, then at what level should the interest rate be set to keep output from changing?

d. Suppose the financial crisis causes $\bar{f}$ to increase as indicated in part (c) and also causes planned autonomous investment to decrease to $\bar{I} = \$1.1$ trillion. Will the change in the interest rate implemented by the Bank of Canada in part (c) be effective in stabilizing output? If not, what additional monetary or fiscal policy changes could be implemented to stabilize output at the original equilibrium output level given in part (b)?

DATA ANALYSIS PROBLEMS

1. Go to the CANSIM database, locate Table 3800067, and find data on household final consumption expenditure (series V61989012), household consumption expenditure on goods (series V61989013), household consumption expenditure on durable goods (series V61989014), household consumption expenditure on semi durable goods (series V61989015), household consumption expenditure on non durable goods (series V61989016), and household consumption expenditure on services (series V61989017).

a. What is the annual growth rate of each series?

b. Which series is more volatile?

2. Go to the CANSIM database, locate Table 228-0069, and find data on Canadian merchandise exports to the United States (V87008956) and Canadian merchandise imports from the United States (V87008840).

a. Move the data into an Excel spreadsheet and calculate the difference between exports and imports. What does this difference represent?

b. Prepare a graph and show the three series, with the exports and imports series on the y_1 axis and the series you calculated in (a) in the y_2 axis. Comment on the values of the series you plotted on the y_2 axis.

22 The Monetary Policy and Aggregate Demand Curves

Learning Objectives

22.1 Recognize the impact of changes in the nominal overnight interest rate on short-term real interest rates.

22.2 Define and illustrate the monetary policy (*MP*) curve, and explain shifts in the *MP* curve.

22.3 Explain why the aggregate demand (*AD*) curve slopes downward, and explain shifts in the *AD* curve.

Preview

At the height of the financial crisis in December 2008, the Federal Open Market Committee of the Federal Reserve announced a surprisingly bold policy decision that sent the markets into a frenzy. The committee lowered the federal funds rate, the interest rate charged on overnight loans between banks, by 75 basis points (0.75 percentage point), moving the federal funds rate almost all the way to zero. In Canada, in response to the global financial crisis, the Bank of Canada also lowered the overnight interest rate from 4.5% in August 2007 to 0.25% in April 2009. Similarly, in response to the Covid-19 pandemic, in March 2020, the Bank of Canada lowered the overnight interest rate from 1.75% to 0.25%, which is what the Bank considers to be its effective lower bound.

To see how monetary policy actions like the ones described above affect the economy, we need to analyze how monetary policy affects aggregate demand. We start this chapter by explaining why monetary policy leads to higher interest rates when inflation increases, leading to a positive relationship between real interest rates and inflation that can be illustrated using the *monetary policy (MP) curve*. Then, by combining the *MP* curve with the *IS* curve we developed in the previous chapter, we derive the *aggregate demand curve*, a key element in the aggregate demand/aggregate supply model framework used in the rest of this book to discuss short-run economic fluctuations.

22.1 THE BANK OF CANADA AND MONETARY POLICY

LO 22.1 Recognize the impact of changes in the nominal overnight interest rate on short-term real interest rates.

Central banks throughout the world use a very short-term interest rate as their primary policy tool. In Canada, the Bank of Canada conducts monetary policy via its setting of the overnight interest rate.

As we saw in Chapter 16, the Bank of Canada controls the overnight interest rate by varying the reserves it provides to the banking system. When it provides more reserves, banks have more money to lend to each other, and this excess liquidity causes the overnight interest rate to fall. When the Bank of Canada drains reserves from the banking system, banks have less to lend, and the lack of liquidity leads to a rise in the overnight interest rate.

The overnight interest rate is a *nominal* interest rate, but as we learned in the previous chapter, it is the *real* interest rate that affects net exports and business spending, thereby determining the level of equilibrium output. How does the Bank of Canada's

control of the overnight interest rate enable it to control the real interest rate, through which monetary policy affects the economy?

Recall from Chapter 4 that the real interest rate, r, is the nominal interest rate, i, minus expected inflation, π^e.

$$r = i - \pi^e$$

Changes in nominal interest rates can change the real interest rate only if actual and expected inflation do not change much in the short run. Because prices typically are slow to move, changes in monetary policy will not have an immediate effect on inflation and expected inflation in the short run. As a result, ***when the Bank of Canada lowers the overnight interest rate, real interest rates fall; when the Bank of Canada raises the overnight interest rate, real interest rates rise***.

22.2 THE MONETARY POLICY CURVE

LO 22.2 Define and illustrate the monetary policy (*MP*) curve, and explain shifts in the *MP* curve.

We have seen how the Bank of Canada can control real interest rates in the short run. The next step in our analysis is to examine how monetary policy reacts to inflation. The **monetary policy (*MP*) curve** indicates the relationship between the real interest rate set by the central bank and the inflation rate. We can write the equation of this curve as follows:

$$r = \bar{r} + \lambda\pi \tag{1}$$

where $\bar{r}$ is the autonomous (exogenous) component of the real interest rate set by the monetary policy authorities, which is unrelated to the current level of the inflation rate or any other variable in the model, and λ is the responsiveness of the real interest rate to the inflation rate.

To make our discussion of the monetary policy curve more concrete, Figure 22-1 shows an example of a monetary policy curve *MP* in which $\bar{r} = 1.0$ and $\lambda = 0.5$:

$$r = 1.0 + 0.5\pi \tag{2}$$

At point A, where inflation is 1%, the Bank of Canada sets the real interest rate at 1.5%; at point B, where inflation is 2%, the Bank sets the real interest rate at 2%; and at point C, where inflation is 3%, the Bank sets the real interest rate at 2.5%. The line going through points A, B, and C is the monetary policy curve *MP*, and it is upward-sloping, indicating that monetary policymakers raise real interest rates when the inflation rate rises.

Why the Monetary Policy Curve Has an Upward Slope

The simplest way to see why the *MP* curve has an upward slope is to recognize that central banks seek to keep inflation stable. To stabilize inflation, monetary policymakers tend to follow the **Taylor principle**, named after John Taylor of Stanford University: They raise real interest rates when there is a rise in inflation, as illustrated by the *MP* curve.[1] Central banks raise real interest rates when inflation increases because the rise in the real interest rate would be needed to contract the economy in order to bring inflation back down.

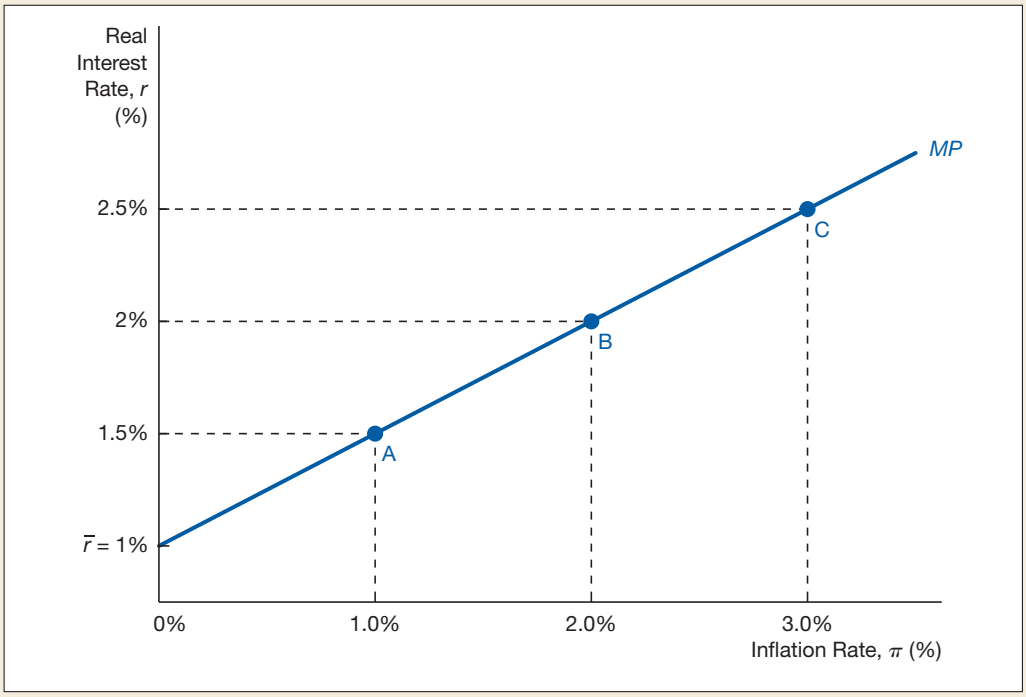
[1]Note that the Taylor principle differs from the Taylor rule, described in Chapter 17, in that it does not provide a rule for how monetary policy should react to conditions in the economy, whereas the Taylor rule does.

Mini-lecture

FIGURE 22-1

The Monetary Policy Curve

The upward slope of the *MP* curve indicates that an outcome from monetary policy is that a rise in inflation leads to a rise in the real interest rate.



An alternative rationale for why monetary policy leads to an upward-sloping *MP* curve results from the central bank not increasing liquidity in the banking system when there is a rise in inflation. Firms and households hold money to carry out transactions. When the inflation rate rises, the price level is higher than it otherwise would be, so more money is needed to pay for the goods and services they want to buy. If the central bank does not increase liquidity in the banking system and create more money for households and firms when inflation rises, households and firms would try to increase their money holdings by moving funds from interest-bearing assets into money. For example, they might increase their money holdings by selling bonds, which would drive down bond prices and therefore raise interest rates. Alternatively, they would try to increase their money holdings by reducing their interest-bearing savings accounts at banks, which would then have less funds to lend out, which would also drive interest rates up. The resulting rise in the nominal interest rate when the monetary policy does not increase liquidity in the banking system would raise the real interest rate, because in the short run, inflation expectations are unchanged, so nominal and real interest rates move together.

Shifts in the *MP* Curve

In common parlance, the Bank of Canada is said to "tighten" monetary policy when it raises real interest rates and to "ease" it when it lowers real interest rates. It is important, however, to distinguish between changes in monetary policy that shift the monetary policy curve, which we call *autonomous* changes, and the Taylor principle–driven changes that are reflected in movements along the monetary policy curve, which are called *automatic* adjustments to interest rates.

Central banks may make autonomous changes to monetary policy for various reasons. They may wish to change the inflation rate from its current value. For example, to lower inflation, they could increase $\bar{r}$ by one percentage point and so raise the real interest rate at any given inflation rate, a move that we will refer to as an **autonomous tightening of monetary policy**. At a 2% inflation rate, the real interest rate would rise from 2% to, say, 3%—that is, from point B to point B_2 in Figure 22-2. This autonomous monetary tightening would shift the monetary policy curve upward by one percentage point, from MP_1 to MP_2, thereby causing the economy to contract and inflation to fall. Or, the central banks may have other information, unrelated to inflation, suggesting that interest rates must be adjusted to achieve good economic outcomes. For example, if the economy is going into a recession, monetary policymakers will want to lower real interest rates at any given inflation rate, an **autonomous easing of monetary policy**, in order to stimulate the economy and to prevent inflation from falling. In this case, at a 2% inflation rate, the real interest rate would fall from 2% to, say, 1%—that is, from point B to point B_3. This autonomous easing of monetary policy would result in a downward shift of the monetary policy curve by, say, one percentage point, from MP_1 to MP_3 in Figure 22-2.

Movements Along Versus Shifts in the *MP* Curve

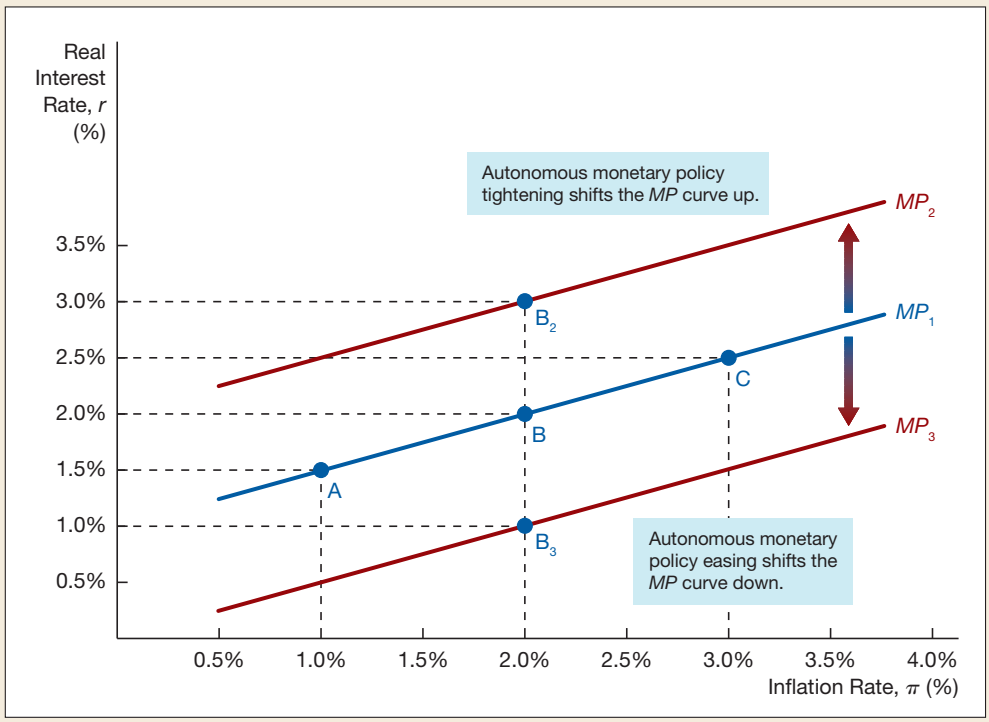
A stumbling block for many students studying the aggregate demand/aggregate supply (*AD/AS*) framework is understanding the distinction between *shifts in* the *MP* curve and *movements along* the *MP* curve. Movements along the *MP* curve—that is, movements from point A to B to C in Figure 22-1—should be viewed as a central bank's normal response (also known as an endogenous response) of raising interest rates when inflation is rising. Thus we can think of a movement along the *MP* curve as an *automatic* response of the central bank to a change in inflation. Such an automatic response does not involve a shift in the *MP* curve.

Mini-lecture

FIGURE 22-2

Shifts in the Monetary Policy Curve

Autonomous changes in monetary policy—for example, when a central bank changes the real interest rate at any given inflation rate—shift the *MP* curve. An autonomous tightening of monetary policy that increases the real interest rate shifts the *MP* curve up to MP_2, whereas an autonomous easing of monetary policy that lowers the real interest rate shifts the *MP* curve down to MP_3.



However, when the central bank raises interest rates *at a given level of the inflation rate*, this action is not an automatic response to higher inflation but is instead an autonomous tightening of monetary policy that shifts the *MP* curve up, from MP_1 to MP_2 in Figure 22-2.

The distinction between autonomous monetary policy changes and movements along the monetary policy curve is illustrated by the two following applications, which describe the monetary policy actions taken by the Federal Reserve in the period from 2004 to 2006 and at the onset of the global financial crisis in the fall of 2007.

APPLICATION Movements Along the *MP* Curve: The Rise in the Federal Funds Rate Target, 2004–2006 and 2015–2019

Fears of deflation—that is, fears that inflation could turn negative—led the Federal Reserve to commit to the very low federal funds rate of 1% from June 2003 to June 2004. However, with the economy growing rapidly, inflationary pressures began to rise, and at its June 2004 meeting the FOMC decided to increase the federal funds rate by ¼ of a percentage point. Furthermore, the FOMC made this increase a very automatic process, raising the federal funds rate by exactly the same amount at every subsequent FOMC meeting through June 2006 (see Figure 22-3). Then, from 2015 to 2019, inflation began

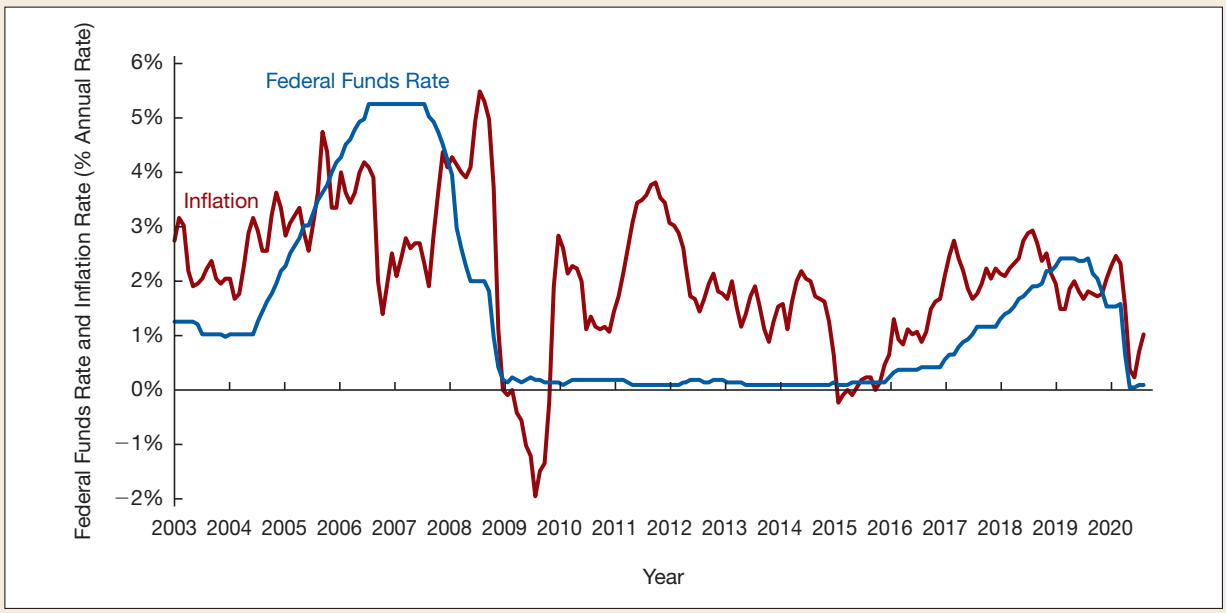


FIGURE 22-3 The Federal Funds Rate and Inflation Rate, 2003–2020

From June 2004 through June 2006, because of pressures from rising inflation, the Fed increased its policy rate, the federal funds rate, by ¼ of a percentage point at every FOMC meeting. Then, from 2015 to 2019, rises in inflation led to the Fed raising the federal funds rate from zero to 2%. These Fed actions were a movement along the *MP* curve. In contrast, the Fed began an autonomous easing of monetary policy in September 2007 and in March 2020, bringing down the federal funds rate sharply, even though inflation at the time was not falling.

Source: Federal Reserve Bank of St. Louis FRED database: http://research.stlouisfed.org/fred2/series/CPIAUCSL; http://research.stlouisfed.org/fred2/series/FEDFUNDS.

to rise, and the FOMC raised the federal funds rate starting in December of 2015 until it reached a level of 2% in 2019. What do these episodes tell us about the monetary policy curve during these time periods?

Because the Federal Reserve was reacting to inflationary pressures in both periods, its monetary policy actions were clearly movements along the *MP* curve, say, from point A to B to C in Figure 22-1. The Fed was following the Taylor principle, reacting to higher inflation by raising the real interest rate. ◆

APPLICATION Shift in the *MP* Curve: Autonomous Monetary Easing in Canada During the Global Financial and Coronavirus Crises

When the global financial crisis started in August 2007, inflation was rising and the unemployment rate was quite low, yet the Bank of Canada began an aggressive easing of monetary policy, lowering the overnight interest rate as shown in Figure 22-4. In March

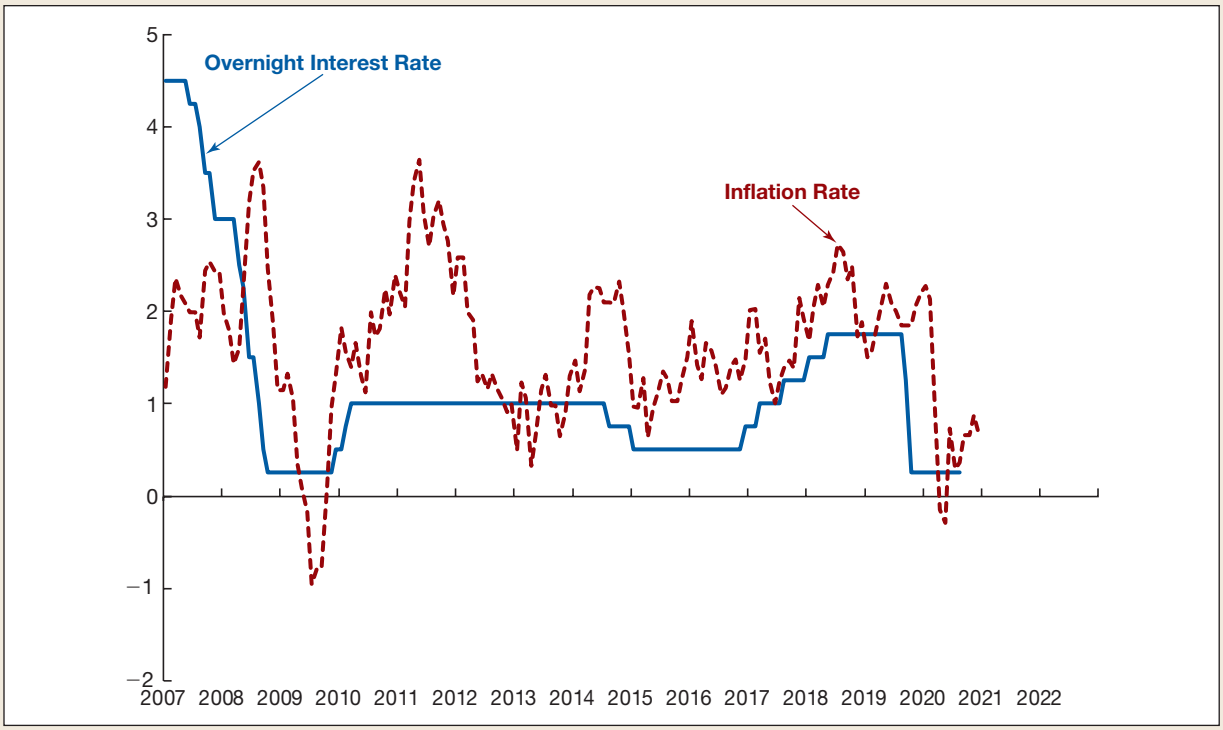


FIGURE 22-4 The Canadian Overnight Interest Rate and Inflation Rate, 2007–2020

The Bank of Canada began an autonomous easing of monetary policy in September 2007, and again in March 2020, bringing down the overnight interest rate despite the continuing high inflation.

Source: Statistics Canada CANSIM series V41690914 and V39079.

2020, the inflation rate was near its target level of 2%, and the unemployment rate was again very low. When it became clear that the coronavirus was beginning to spread out of control, the Bank of Canada decided to lower the overnight interest rate by 50 basis points (0.50 percentage point) on March 4, 2020, by another 50 basis points on March 13, 2020, and then by another 50 basis points on March 27, 2020, down to 0.25%, which is what the Bank considers to be its effective lower bound. What were the implications for the monetary policy curve?

A movement along the *MP* curve would have suggested that the Bank of Canada planned to continue hiking interest rates because inflation was rising, but instead the Bank of Canada did the opposite. Similarly, with unemployment low and the inflation rate near the target level, when the coronavirus crisis was first beginning, the Taylor principle would not have suggested that the Bank of Canada would lower interest rates, although it did do so dramatically in March. The Bank of Canada's actions in both cases thus shifted the monetary policy curve down, from MP_1 to MP_3, as shown in Figure 22-2. The Bank of Canada pursued this autonomous monetary policy easing because the negative shock to the economy caused by the disruption to financial markets and the coronavirus pandemic (discussed in Chapter 11) indicated that the economy was likely to weaken in the near future, and the inflation rate would then fall. Indeed, this is exactly what came to pass, with the economy going into a recession after the onset of the global financial and coronavirus crises. ◆

22.3 THE AGGREGATE DEMAND CURVE

LO 22.3 Explain why the aggregate demand (*AD*) curve slopes downward, and explain shifts in the *AD* curve.

We are now ready to derive the relationship between the inflation rate and aggregate output when the goods market is in equilibrium: the **aggregate demand curve**. The *MP* curve that we just developed demonstrates how central banks respond to changes in inflation by changing the interest rate in line with the Taylor principle. The *IS* curve we developed in Chapter 21 showed that changes in real interest rates, in turn, affect equilibrium output. With these two curves, we can now link the quantity of aggregate output demanded with the inflation rate, given the public's expectations of inflation and the stance of monetary policy. The aggregate demand curve is central to the aggregate demand and supply analysis we will develop further in the next chapter, which will enable us to explain short-run fluctuations in both aggregate output and inflation.

Deriving the Aggregate Demand Curve Graphically

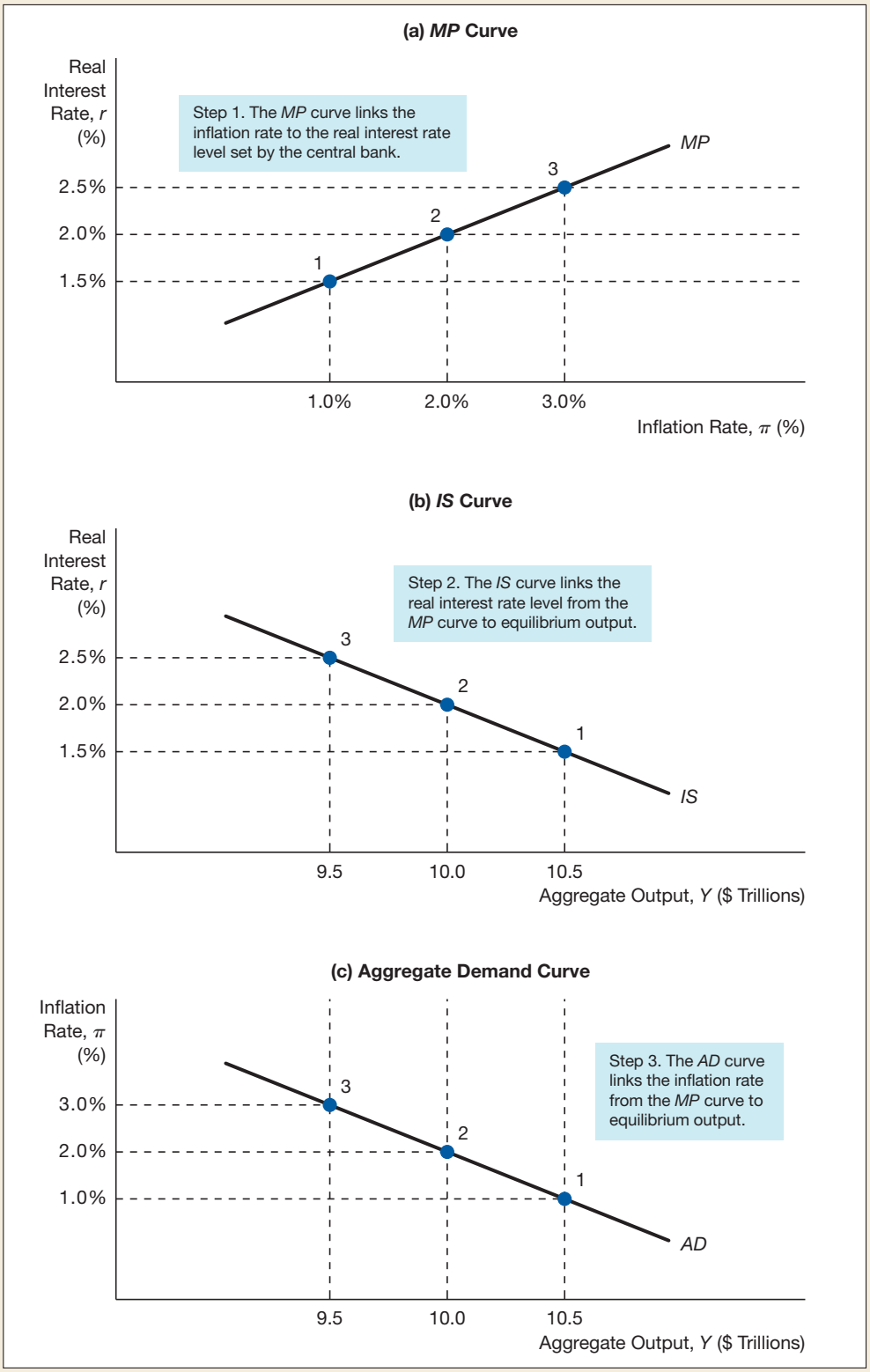
Using the hypothetical *MP* curve from Equation 2, we know that when the inflation rate rises from 1% to 2% to 3%, the Bank of Canada reacts by raising the real interest rate from 1.5% to 2% to 2.5%. We plot these points in panel (a) of Figure 22-5 as the *MP* curve. In panel (b), we graph the *IS* curve described in Equation 13 of Chapter 21 ($Y = 12 - r$). As the real interest rate rises from 1.5% to 2% to 2.5%, the equilibrium

Mini-lecture

FIGURE 22-5

Deriving the *AD* Curve

The *MP* curve in panel (a) shows that as inflation rises from 1.0% to 2.0% to 3.0%, the real interest rate rises from 1.5% to 2.0% to 2.5%. The *IS* curve in panel (b) then shows that higher real interest rates lead to lower planned investment spending and net exports, and hence aggregate output falls from $10.5 trillion to $10.0 trillion to $9.5 trillion. Finally, panel (c) plots the level of equilibrium output corresponding to each of the three inflation rates: The line that connects these points is the *AD* curve, and it is downward-sloping.



moves from point 1 to point 2 to point 3, and aggregate output falls from $10.5 trillion to $10 trillion to $9.5 trillion. In other words, as real interest rates rise, investment and net exports decline, leading to a reduction in aggregate demand. Using the information from panels (a) and (b), we can create the curve shown in panel (c). As inflation rises from 1% to 2% to 3%, the equilibrium moves from point 1 to point 2 to point 3 in panel (c), and aggregate output falls from $10.5 trillion to $10 trillion to $9.5 trillion.

The line that connects the three points shown in panel (c) is the aggregate demand curve, *AD,* and it indicates the level of aggregate output corresponding to each of the three real interest rates consistent with equilibrium in the goods market for any given inflation rate. The aggregate demand curve has a downward slope because a higher inflation rate leads the central bank to raise the real interest rate, thereby lowering planned spending and hence lowering the level of equilibrium aggregate output.

By using some algebra (see the FYI box "Deriving the Aggregate Demand Curve Algebraically"), the *AD* curve in Figure 22-5 can be written numerically as follows:

$$Y = 11 - 0.5\pi \tag{3}$$

Factors That Shift the Aggregate Demand Curve

Movements along the aggregate demand curve describe how the equilibrium level of aggregate output changes when the inflation rate changes. When factors other than the inflation rate change, however, the aggregate demand curve shifts. We first review the factors that shift the *IS* curve and then consider the factors that shift the *AD* curve.

FYI Deriving the Aggregate Demand Curve Algebraically

To derive a numerical *AD* curve, we start by taking the numerical *IS* curve given by Equation 13 in Chapter 21:

$$Y = 12 - r$$

We then use the numerical *MP* curve given in Equation 2, $r = 1.0 + 0.5\pi$, to substitute for r, yielding

$$\begin{aligned} Y &= 12 - (1.0 + 0.5\pi) \\ &= (12 - 1) - 0.5\pi \\ &= 11 - 0.5\pi \end{aligned}$$

which is the same as Equation 3 in the text.

Similarly, we can derive a more general version of the *AD* curve by using the algebraic version of the *IS* curve given by Equation 12 in Chapter 21:

$$Y = [\bar{C} + \bar{I} - d\bar{f} + \bar{G} + \overline{NX} - mpc \times \bar{T}] \times \frac{1}{1 - mpc} - \frac{d + x}{1 - mpc} \times r$$

We then substitute for r using the algebraic *MP* curve given in Equation 1, $r = \bar{r} + \lambda\pi$, to get the more general equation of the *AD* curve:

$$Y = [\bar{C} + \bar{I} - d\bar{f} + \bar{G} + \overline{NX} - mpc \times \bar{T}] \times \frac{1}{1 - mpc} - \frac{d + x}{1 - mpc} \times (\bar{r} + \lambda\pi)$$

Shifts in the *IS* Curve We saw in the preceding chapter that there are six factors that cause the *IS* curve to shift. It turns out that these same factors will cause the aggregate demand curve to shift as well:

1. Autonomous consumption expenditure
2. Autonomous investment spending
3. Government purchases
4. Taxes
5. Autonomous net exports
6. Financial frictions

We examine how changes in these factors lead to a shift in the aggregate demand curve, as shown in Figure 22-6.

Suppose that inflation is at 2.0%, and so the *MP* curve in panel (a) of Figure 22-6 shows that the real interest rate is at 2.0%. The IS_1 curve in panel (b) then shows that the equilibrium level of output is at \$10 trillion at point A_1, which corresponds to an equilibrium level of output of \$10 trillion at point A_1 on the AD_1 curve in panel (c). Now suppose there is a rise in, for example, government purchases of \$1 trillion. Panel (b) shows that with both the inflation rate and the real interest rate held constant at 2.0%, the equilibrium moves from point A_1 to point A_2, with output rising to \$12.5 trillion[2] and so the *IS* curve shifts to the right from IS_1 to IS_2. The rise in output to \$12.5 trillion means that, holding inflation and the real interest rate constant, the equilibrium in panel (c) also moves from point A_1 to point A_2, and so the *AD* curve also shifts to the right, from AD_1 to AD_2.

Figure 22-6 shows that ***any factor that shifts the* IS *curve shifts the aggregate demand curve in the same direction***. Therefore, any factor that shifts the *IS* curve to the right—a rise in autonomous consumption expenditure or planned investment spending encouraged by "animal spirits," a rise in government purchases, an autonomous rise in net exports, a fall in taxes, or a decline in financial frictions—will also shift the aggregate demand curve to the right. Conversely, any factor that shifts the *IS* curve to the left—a fall in autonomous consumption expenditure, a fall in planned investment spending, a fall in government purchases, a fall in net exports, a rise in taxes, or a rise in financial frictions—will shift the aggregate demand curve to the left.

Shifts in the *MP* Curve We now examine what happens to the aggregate demand curve when the *MP* curve shifts. Suppose the Bank of Canada is worried about the economy overheating and so decides to autonomously tighten monetary policy by raising the real interest rate by one percentage point at any given level of the inflation rate. At an inflation rate of 2.0%, the real interest rate rises from 2.0% to 3.0% in Figure 22-7. The *MP* curve shifts up from MP_1 to MP_2 in panel (a). Panel (b) shows that when the inflation rate is at 2.0%, the higher real interest rate of 3.0% causes the equilibrium to move from point A_1 to point A_2 on the *IS* curve, with output falling from \$10 trillion to \$9 trillion. The lower output of \$9 trillion occurs because the higher real interest leads to a decline in investment and net exports, which

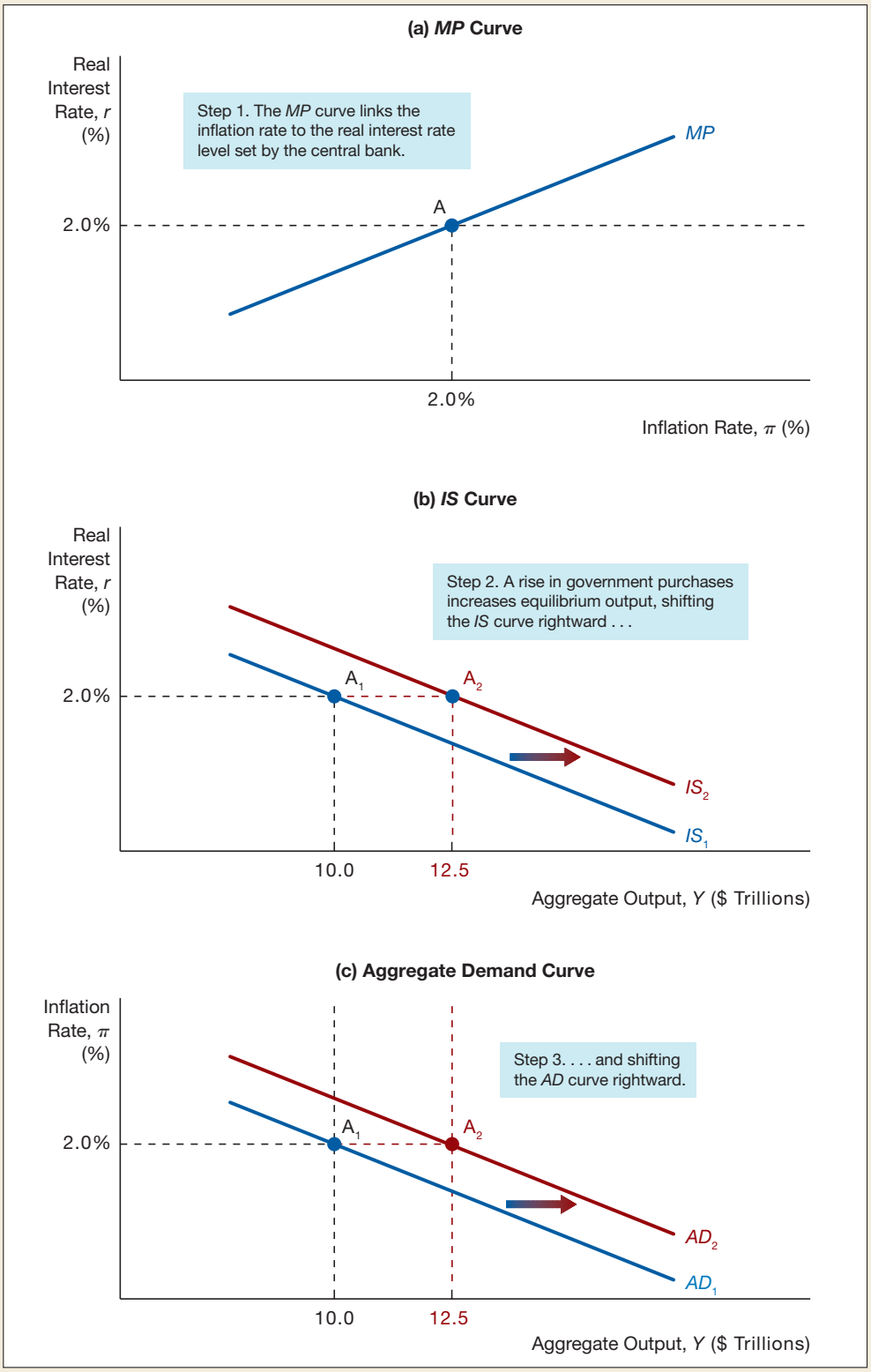
[2]As we saw in the numerical example in Chapter 21, a rise in government purchases of \$1 trillion leads to a \$2.5 trillion increase in equilibrium output at any given real interest rate, and this is why output rises from \$10 trillion to \$12.5 trillion when the real interest rate is at 2.0%.

Mini-lecture

FIGURE 22-6

Shift in the *AD* Curve from Shifts in the *IS* Curve

At a 2% inflation rate in panel (a), the monetary policy curve indicates that the real interest rate is 2%. An increase in government purchases shifts the *IS* curve to the right in panel (b). At a given inflation rate and real interest rate of 2%, equilibrium output rises from $10 trillion to $12.5 trillion, which is shown as a movement from point A_1 to point A_2 in panel (c), shifting the aggregate demand curve to the right from AD_1 to AD_2. Any factor that shifts the *IS* curve shifts the *AD* curve in the same direction.

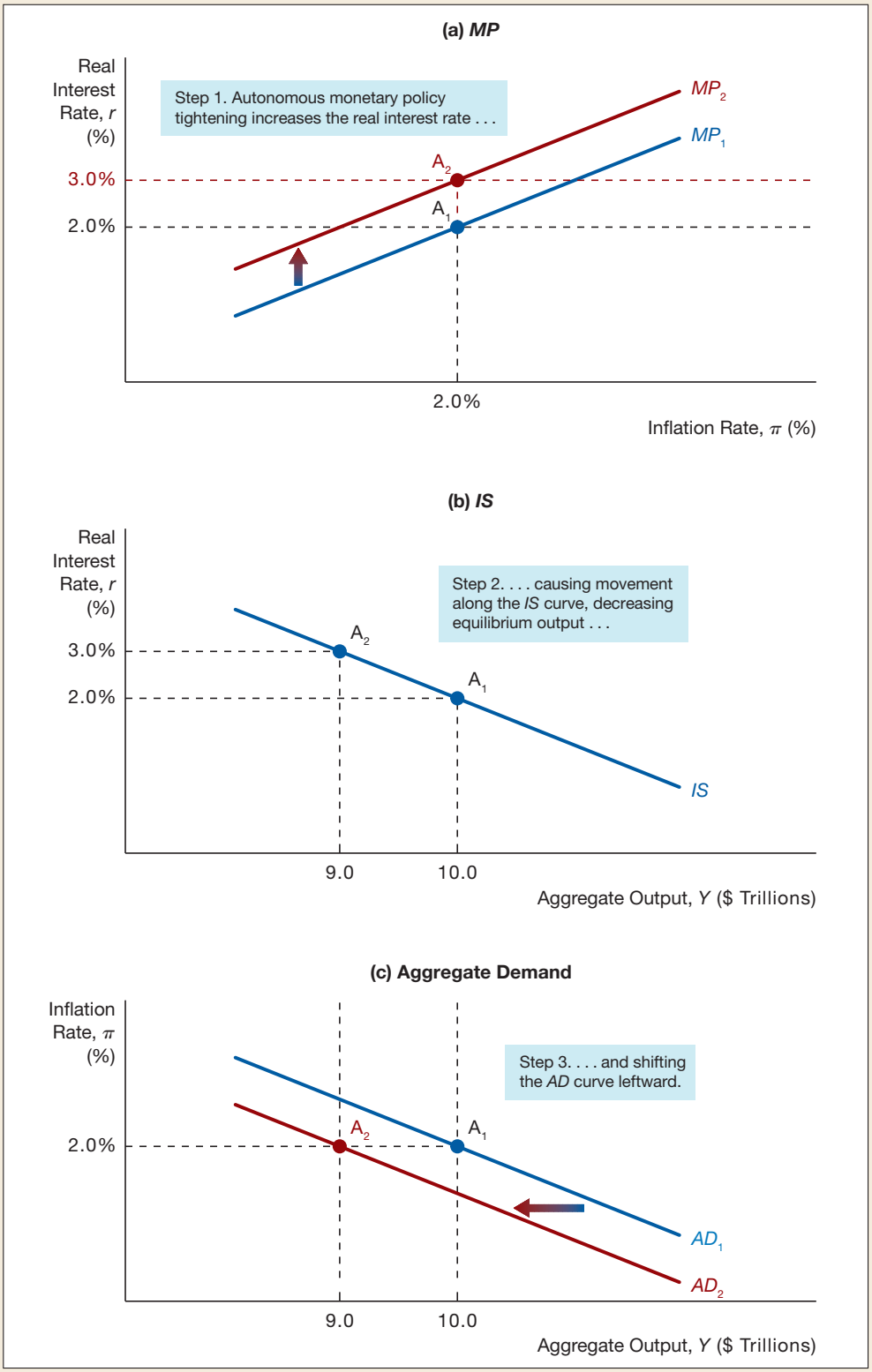


Mini-lecture

FIGURE 22-7

Shift in the *AD* Curve from Autonomous Monetary Policy Tightening

Autonomous monetary tightening that raises real interest rates by one percentage point at any given inflation rate shifts the *MP* curve up, from MP_1 to MP_2 in panel (a). With the inflation rate at 2%, the higher 3% interest rate results in a movement from point A_1 to A_2 on the *IS* curve, with output falling from $10 trillion to $9 trillion. This change in equilibrium output leads to movement from point A_1 to point A_2 in panel (c), shifting the aggregate demand curve to the left, from AD_1 to AD_2.



lowers aggregate demand. The lower output of $9 trillion then decreases the equilibrium output level from point A_1 to point A_2 in panel (c), and so the *AD* curve shifts to the left, from AD_1 to AD_2.

Our conclusion from Figure 22-7 is that ***an autonomous tightening of monetary policy—that is, a rise in the real interest rate at any given inflation rate—shifts the aggregate demand curve to the left. Conversely, an autonomous easing of monetary policy shifts the aggregate demand curve to the right.***

We have now derived and analyzed the aggregate demand curve, an essential element in the aggregate demand and supply framework, that we will examine in the next chapter. We will use the aggregate demand curve in this framework to determine both aggregate output and inflation as well as to examine events that cause these variables to change.

SUMMARY

1. When the Bank of Canada lowers the overnight interest rate by providing more liquidity to the banking system, real interest rates fall in the short run; when the Bank of Canada raises the overnight interest rate by reducing the amount of liquidity in the banking system, real interest rates rise in the short run.
2. The monetary policy (*MP*) curve shows the relationship between inflation and the real interest rate that arises from monetary authorities' actions. Monetary policymakers follow the Taylor principle by raising real interest rates in response to higher inflation, represented by the upward slope of the monetary policy curve. An autonomous tightening of monetary policy occurs when monetary policymakers raise the real interest rate at any given inflation rate, resulting in an upward shift of the monetary policy curve. An autonomous easing of monetary policy and a downward shift of the monetary policy curve occur when monetary policymakers lower the real interest rate at any given inflation rate.
3. The aggregate demand curve gives the level of equilibrium aggregate output (which equals the total quantity of output demanded) for any given inflation rate. It slopes downward because a higher inflation rate leads the central bank to raise real interest rates, which leads to a lower level of equilibrium output. The aggregate demand curve will shift in the same direction as the *IS* curve; hence the *AD* curve shifts to the right when government purchases increase, taxes decrease, "animal spirits" encourage consumer and business spending, autonomous net exports increase, or financial frictions decrease. An autonomous tightening of monetary policy—that is, an increase in the real interest rate at any given inflation rate—leads to a decline in aggregate demand, and the aggregate demand curve shifts to the left.

MyLab Economics Make the grade with MyLab Economics: Select end-of-chapter questions can be found on MyLab Economics. You can practise them as often as you want, and most feature step-by-step guided instructions to help you find the right answer.

QUESTIONS

1. When the inflation rate increases, what happens to the overnight interest rate? Operationally, how does the Bank of Canada adjust the overnight interest rate?
2. What is the key assumption underlying the Bank of Canada's ability to control the real interest rate?
3. Why does the *MP* curve necessarily have an upward slope?
4. If $\lambda = 0$, what does this imply about the relationship between the nominal interest rate and the inflation rate?

5. How does an autonomous tightening or easing of monetary policy by the Bank of Canada affect the *MP* curve?
6. How is an autonomous tightening or easing of monetary policy different from a change in the real interest rate caused by a change in the current inflation rate?
7. Suppose that a new Bank of Canada governor is appointed and that their approach to monetary policy can be summarized by the following statement: "I care only about increasing employment. Inflation has been at very low levels for quite some time; my priority is to ease monetary policy to promote employment." How would you expect the monetary policy curve to be affected, if at all?
8. "The Bank of Canada decreased the overnight interest rate in late 2007, even though inflation was increasing. This action demonstrated a violation of the Taylor principle." Is this statement true, false, or uncertain? Explain your answer.
9. What factors affect the slope of the aggregate demand curve?
10. "Autonomous monetary policy is more effective at changing output when λ is higher." Is this statement true, false, or uncertain? Explain your answer.
11. If net exports were not sensitive to changes in the real interest rate, would monetary policy be more or less effective in changing output?
12. If an asset-price bubble begins to form, assuming the central bank responds, how is it likely to respond, and what will be the effect on the *MP* curve?
13. For each of the following situations, describe how (if at all) the *IS*, *MP*, and *AD* curves are affected.
    a. A decrease in financial frictions
    b. An increase in taxes and an autonomous easing of monetary policy
    c. An increase in the current inflation rate
    d. A decrease in autonomous consumption
    e. Firms become more optimistic about the future of the economy.
    f. The new Bank of Canada governor begins to care more about fighting inflation.
14. What would be the effect of an increase in Canadian net exports on the aggregate demand curve? Would an increase in net exports affect the monetary policy curve? Explain.
15. Why does the aggregate demand curve shift when "animal spirits" change?
16. Suppose that government spending is increased at the same time that an autonomous monetary policy tightening occurs. What will happen to the position of the aggregate demand curve?
17. If financial frictions increase, how will this affect credit spreads, and how might the central bank respond? Why?
18. "If $\bar{f}$ increases, then the Bank of Canada can keep output constant by reducing the real interest rate by the same amount as the increase in financial frictions." Is this statement true, false, or uncertain? Explain your answer.

APPLIED PROBLEMS

19. Assume that the monetary policy curve is given by $r = 1.5 + 0.75\pi$.
    a. Calculate the real interest rate when the inflation rate is 2%, 3%, and 4%.
    b. Draw a graph of the *MP* curve, labelling the points from part (a).
    c. Assume now that the monetary policy curve is given by $r = 2.5 + 0.75\pi$. Does the new monetary policy curve represent an autonomous tightening or loosening of monetary policy?
    d. Calculate the real interest rate when the inflation rate is 2%, 3%, and 4%, and draw the new *MP* curve, showing the shift from part (b).
20. Use an *IS* curve and an *MP* curve to derive graphically the *AD* curve.
21. Suppose the monetary policy curve is given by $r = 1.5 + 0.75\pi$, and the *IS* curve is given by $Y = 13 - r$.
    a. Calculate an expression for the aggregate demand curve.
    b. Calculate the real interest rate and aggregate output when the inflation rate is 2%, 3%, and 4%.
    c. Draw graphs of the *IS*, *MP*, and *AD* curves, labelling the points from part (b) on the appropriate graphs.
22. Consider an economy described by the following:
    $\bar{C} =$ \$4 trillion
    $\bar{I} =$ \$1.5 trillion
    $\bar{G} =$ \$3.0 trillion
    $\bar{T} =$ \$3.0 trillion

$\overline{NX}$ = \$1.0 trillion
$\bar{f}$ = 0
mpc = 0.8
d = 0.35
x = 0.15
λ = 0.5
$\bar{r}$ = 2

a. Derive expressions for the *MP* curve and the *AD* curve.
b. Calculate the real interest rate and aggregate output when $\pi = 2$ and $\pi = 4$.
c. Draw a graph of the *MP* curve and the *AD* curve, labelling the points given in part (b).

23. Consider an economy described by the following:

$\overline{C}$ = \$3.25 trillion
$\bar{I}$ = \$1.3 trillion
$\overline{G}$ = \$3.5 trillion
$\overline{T}$ = \$3.0 trillion
$\overline{NX}$ = −\$1.0 trillion
$\bar{f}$ = 1
mpc = 0.75
d = 0.3
x = 0.1
λ = 1
$\bar{r}$ = 1

a. Derive expressions for the *MP* curve and the *AD* curve.
b. Assume that $\pi = 1$. Calculate the real interest rate, the equilibrium level of output, consumption, planned investment, and net exports.
c. Suppose the Bank of Canada increases $\bar{r}$ to $\bar{r} = 2$. Calculate the real interest rate, the equilibrium level of output, consumption, planned investment, and net exports at this new level of $\bar{r}$.
d. Considering that output, consumption, planned investment, and net exports all decreased in part (c), why might the Bank of Canada choose to increase $\bar{r}$?

24. Consider the economy described in Applied Problem 23.

a. Derive expressions for the *MP* curve and the *AD* curve.
b. Assume that $\pi = 2$. What are the real interest rate and the equilibrium level of output?
c. Suppose government spending increases to \$4 trillion. What happens to equilibrium output?
d. If the Bank of Canada wants to keep output constant, what monetary policy change should it make?

25. Suppose the *MP* curve is given by $r = 2 + \pi$, and the *IS* curve is given by $Y = 20 - 2r$.

a. Derive an expression for the *AD* curve, and draw a graph labelling points at $\pi = 0$, $\pi = 4$, and $\pi = 8$.
b. Suppose that λ increases to $\lambda = 2$. Derive an expression for the new *AD* curve, and draw the new *AD* curve using the graph from part (a).
c. What does your answer to part (b) imply about the relationship between a central bank's distaste for inflation and the slope of the *AD* curve?

DATA ANALYSIS PROBLEMS

1. A measure of real interest rates can be approximated by the interest rate on inflation-protected long-term Canada bonds, or Real Return Bonds. Go to the CANSIM database and find monthly data from January 1992 to December 2020 on the long-term Real Return Bond (series V122553) and the (all items) consumer price index (CPI), a measure of the price index (series V41690914). Convert the price index data to annualized inflation rates by taking the month-to-month percent change in the price index and multiplying it by 12. Be sure to multiply by 100 so that your results are percentages.

a. Calculate the average inflation rate and the average real interest rate over the most recent 12 months of data available.
b. Calculate the standard deviation of the real interest rate and the annualized inflation rate over the most recent 12 months of data available.

2. Continuing from the previous problem, use the monthly CPI and calculate the percent change from a year ago. Then plot both series on the same graph, using data from 2007 through the most current data available.

23 Aggregate Demand and Supply Analysis

Learning Objectives

23.1 Describe the variables that aggregate demand and supply analysis are trying to explain.

23.2 Summarize and illustrate the aggregate demand curve and the factors that shift it.

23.3 Illustrate and interpret the short-run and long-run aggregate supply curves.

23.4 Illustrate and interpret shifts in the short-run and long-run aggregate supply curves.

23.5 Illustrate and interpret the short-run and long-run equilibria and the role of the self-correcting mechanism.

23.6 Illustrate and interpret the short-run and long-run effects of a shock to aggregate demand.

23.7 Illustrate and interpret the short-run and long-run effects of supply shocks.

23.8 Summarize the conclusions from aggregate demand and supply analysis.

Preview

In earlier chapters, we focused considerable attention on monetary policy because it touches our everyday lives by affecting both inflation and the quantity of jobs available. In this chapter, we develop aggregate demand and supply analysis, a basic tool that will enable us to explain the effects of monetary policy on fluctuations in inflation, aggregate output, and unemployment. (The Following the Financial News box "Aggregate Output, Unemployment, and Inflation" indicates where and how often data on aggregate output, unemployment, and the inflation rate are published.) Not only will the analysis help us interpret the past; it also will help us understand recent episodes in the business cycle, such as the severe recessions resulting from the global financial crisis of 2007–2009 and the coronavirus pandemic in 2020. It also will help predict how future events may affect aggregate output, unemployment, and inflation.

23.1 BUSINESS CYCLES AND INFLATION

LO 23.1 Describe the variables that aggregate demand and supply analysis are trying to explain.

Before we develop aggregate demand and supply analysis, let's look at what this analysis is trying to explain: business cycles and inflation. First, let's discuss business cycles.

Business Cycles

Business cycles are the upward and downward movements of aggregate economic activity produced in the economy. Real GDP, the total of goods and services produced in a year, using constant prices, is the most widely used measure of aggregate economic activity, also referred to as **aggregate output**. Many economic variables move up and down together during business cycles. Most importantly, as Figure 23-1 shows, employment rises when real GDP is rising, and employment falls when real GDP falls. **Recessions**, the shaded areas in the figure, are periods when aggregate economic activity is decreasing or contracting. **Expansions** (also called **booms**) are periods when aggregate economic activity is increasing or expanding.

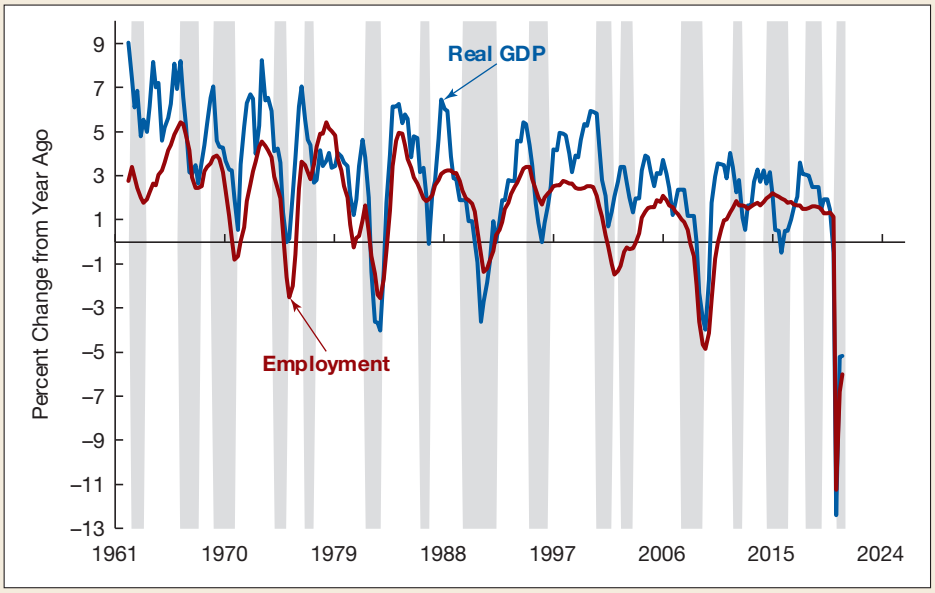
One way to look at business cycles is to divide aggregate output (real GDP) into two components, the long-run trend, which grows smoothly over time, and fluctuations

FIGURE 23-1

GDP and Employment Growth, 1961–2020

The figure shows the growth rate of real GDP and employment, with the recession periods (peak to trough) shaded. Note how closely real GDP and employment growth move together. (*Note:* Shaded areas indicate recessions.)

Source: Federal Reserve Bank of St. Louis, FRED database: https://fred.stlouisfed.org (series LFEMTTTTCAM647S) and Statistics Canada CANSIM series V62305752.



around this trend. The long-run trend of aggregate output, which grows smoothly over time, is referred to as **potential output** (or **potential GDP**). It is the level of aggregate output that is generated when the economy is at full employment and firms on average are producing at their long-run, profit-maximizing level.

We can think about the business cycle as a series of short-run deviations of aggregate output from its long-run trend (that is, from potential output). The **output gap** is the percentage difference between actual aggregate output and potential output.[1] The output gap is measured as the *percentage* difference between actual real GDP and potential GDP. For example, in the second quarter of 2020, the output gap was –10.2%, indicating that aggregate output was 10.2% *below* potential output. Figure 23-2 shows how the output gap has fluctuated since 1961, with the shaded areas indicating recessions. The output gap allows us to compare the magnitude of different business cycle fluctuations. For example, it shows that the 1981–1982 recession, the 2007–2009 recession, and the 2020 Covid-19 recession had output gaps that were large and negative, with real GDP falling significantly below potential GDP.

[1]The output gap is calculated as the difference between real GDP and potential GDP, divided by potential GDP, which is then expressed as a percentage by multiplying by 100.

In the second quarter of 2020, real GDP was $1.84 trillion, while potential GDP (based on the Hodrick and Prescott filter) was $2.05 trillion. The calculation for the output gap for that quarter is

$$\text{Output Gap} = \frac{\text{Real GDP} - \text{Potential GDP}}{\text{Potential GDP}} = \frac{1.84 - 2.05}{2.05} = -0.102 = -10.2\%$$

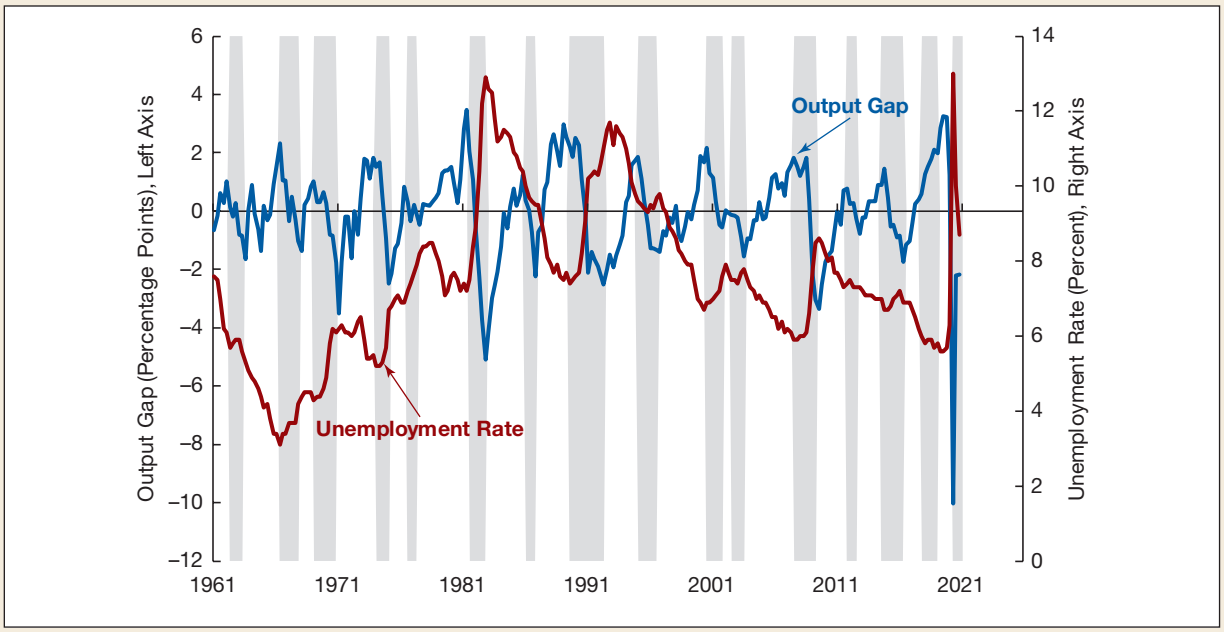


FIGURE 23-2 Output Gap and Unemployment Rate, 1961–2020

The output gap is a measure of the magnitude of business cycle fluctuations, and it shows that the three most severe recessions since 1960 were the 1981–1982 recession, the 2007–2009 recession, and the 2020 Covid-19 recession. The figure shows that the output gap is strongly negatively related to the unemployment rate so that when the output gap falls, the unemployment rate rises, and vice versa. (*Note:* Shaded areas indicate recessions.)

Source: Statistics Canada CANSIM series V62305752 and V20628159 and authors' calculations.

One of the key reasons we care about business cycles is also shown in Figure 23-2, which plots not only business cycle fluctuations as represented by the output gap but also the **unemployment rate**, the percentage of workers looking for work but who do not have jobs. As Figure 23-2 shows, there is a strong negative relationship between the unemployment rate and the output gap. During booms, aggregate output is rising relative to potential output and the output gap rises. (Note that as you can see in the figure, if output is below potential output so the output gap is negative, when the output gap rises, then the output gap becomes less negative. When output is above potential output, then the output gap becomes more positive as it rises.) During business cycle booms when the output gap is rising, firms hire more workers, so the unemployment rate declines.

Clearly, business cycle fluctuations produce large swings in the unemployment rate, with recessions hitting workers very hard when the unemployment rate rises sharply. When unemployment is high, it is more likely that you might lose your job. Furthermore, if you are unlucky enough to graduate when unemployment is high, you will have a hard time finding the dream job you were anticipating.

When people lose their jobs during a recession, many of them can no longer afford to make their mortgage payments and end up losing their homes. Poverty rates also rise. Many people also become depressed during recessions because their self-worth is damaged when they are unemployed and cannot find a job. Their health can further deteriorate if they are unable to pay for adequate health care.

Just as workers do not fare well during recessions, neither do businesses, which often find that (1) they cannot sell all the goods they have produced and (2) they receive lower prices for their products. During recessions, business profits decline. Many firms now have to close their factories or can no longer pay back their loans and end up declaring bankruptcy. Many businesses close their doors, never to open them again.

We care about business cycles because they lead to large fluctuations in unemployment and output. Booms are good for workers and businesses, while recessions lead to great economic hardship. We need to understand what drives business cycles, and as we will see, aggregate demand and supply analysis enables us to do exactly that. Indeed, we will use this analysis to explain major business cycle episodes such as the 1981–1982 recession, the recession of 2007–2009, and the Covid-19 recession that started in 2020.

Inflation

We also care about inflation because low and stable inflation reduces the need for businesses to change prices frequently, which otherwise would be costly. Low and stable inflation also lowers uncertainty in the economy, which makes it easier for households and businesses to make economic decisions, thus increasing the efficiency of the economy.

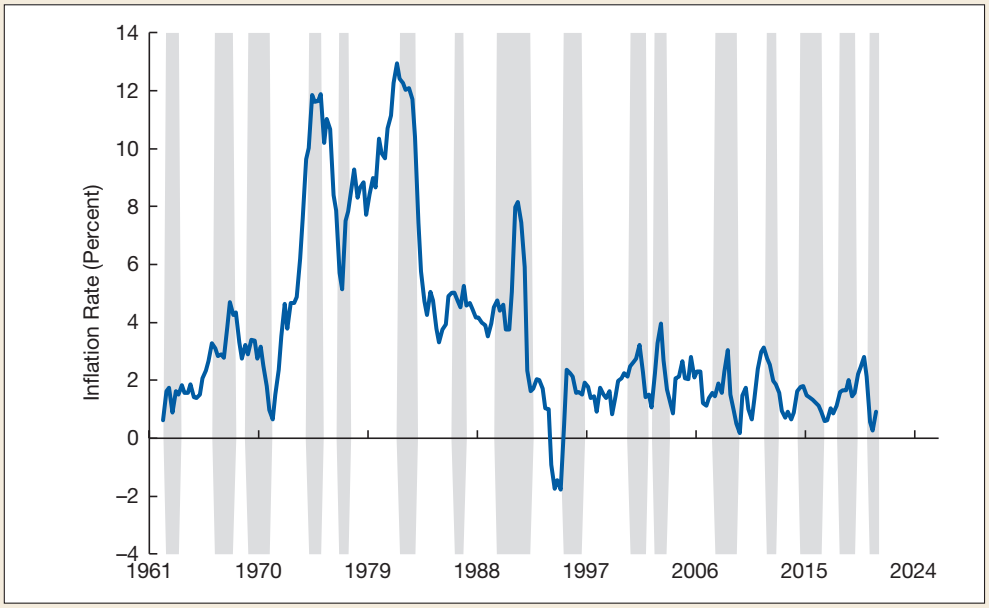
Figure 23-3 shows that inflation has not always been low and stable. For example, the inflation rate reached a peak of about 13% in 1981. In contrast, the inflation rate fell precipitously during the 2007–2009 recession. We will also use aggregate demand and supply analysis to explain fluctuations in inflation, such as those that occurred in these two episodes.

FIGURE 23-3

Inflation Rate, 1961–2020

The inflation rate has not always been low and stable. In 1980, it peaked at about 13%, and in 1994, it even went negative for a brief period of time. (*Note:* Shaded areas indicate recessions.)

Source: Federal Reserve Bank of St. Louis, FRED database: https://fred.stlouisfed.org (series CPALCY01CAM661N).



23.2 AGGREGATE DEMAND

LO 23.2 Summarize and illustrate the aggregate demand curve and the factors that shift it.

We start our analysis of aggregate demand by examining the components of aggregate demand.

Components of Aggregate Demand

Aggregate quantity demanded is the total quantity of goods and services that households, businesses, governments, and foreigners want to buy (demand) at any given inflation rate during a specific period, such as a year. To understand what determines the aggregate quantity demanded, we examine its components.

Aggregate quantity demanded is made up of four component parts: **consumption expenditure**, the total demand for consumer goods and services; **planned investment spending**,[2] the total planned spending by business firms on new machines, factories, and other capital goods, plus planned spending on new homes; **government purchases**, spending by all levels of government (federal, provincial, and local)

Following the Financial News — Aggregate Output, Unemployment, and Inflation

Newspapers and websites periodically report data that provide information on the level of aggregate output, unemployment, and the price level. Here is a list of the relevant data series, their frequency, and when they are published.

Aggregate Output and Unemployment

Real GDP: Quarterly, published about two months after the end of a quarter.

Industrial production: Monthly. Industrial production is not as comprehensive a measure of aggregate output as real GDP because it measures only manufacturing output: The estimate for the previous month is reported after about two months.

Unemployment rate: Monthly. The estimate for a month is reported two months later.

Price Level

GDP deflator: Quarterly. This comprehensive measure of the price level (described in the appendix to Chapter 1) is published at the same time as the real GDP data.

Consumer price index (CPI): Monthly. The CPI is a measure of the price level for consumers (also described in the appendix to Chapter 1); the value for the previous month is published in the third or fourth week of the following month.

Producer price index (PPI): Monthly. The PPI is a measure of the average level of wholesale prices charged by producers and is published in the third or fourth week of the following month.

[2]Recall that economists restrict the definition of the word *investment* to mean the purchase of new physical capital, such as a new machine or a new house, which adds to spending on newly produced goods or services. This differs from the everyday use of the term by noneconomists, who use the word *investment* to describe purchases of common stocks or bonds, purchases that do not necessarily involve newly produced goods and services. When economists speak of investment spending, they are referring to purchases that add to aggregate demand.

on goods and services (paper clips, computers, computer programming, missiles, government employees, and so on); and **net exports**, the net foreign spending on domestic goods and services, equal to exports minus imports. Using the symbols C for consumption expenditure, I for planned investment spending, G for government spending, and NX for net exports, we can write the following expression for aggregate demand Y^{ad}:

$$Y^{ad} = C + I + G + NX \tag{1}$$

Deriving the Aggregate Demand Curve

The first building block of aggregate demand and supply analysis is the **aggregate demand curve**, which describes the relationship between the aggregate quantity demanded and the inflation rate when all other variables are held constant.

The first step in deriving the aggregate demand curve is to recognize that when the inflation rate rises ($\pi \uparrow$), the real interest rate rises ($r \uparrow$). This occurs for either of two reasons: When the inflation rate rises, (1) the monetary authorities take actions to raise the real interest rate to keep inflation from spiralling out of control or (2) the monetary authorities just do not increase liquidity in the banking system when inflation rises, so the real interest rate rises.

Next we can examine the effects of the higher real interest rate on the individual components of aggregate demand. When the real interest rate is higher, the cost of financing purchases of new physical capital becomes higher, making investment less profitable and causing planned investment spending to decline ($I \downarrow$). Because, as shown in Equation 1, planned investment spending is included in aggregate demand, the decline in planned investment spending causes aggregate demand to fall ($Y^{ad} \downarrow$). A higher inflation rate therefore leads to a lower level of the quantity of aggregate output demanded ($\pi \uparrow \implies Y^{ad} \downarrow$), and so the aggregate demand curve slopes downward, as in Figure 23-4.

Schematically, we can describe why the aggregate demand curve slopes downward, as follows:[3]

$$\pi \uparrow \Rightarrow r \uparrow \Rightarrow I \downarrow \Rightarrow Y^{ad} \downarrow$$

Factors That Shift the Aggregate Demand Curve

Seven basic factors (often referred to as **demand shocks**) can shift the aggregate demand curve to a new position: (1) autonomous monetary policy, (2) government purchases, (3) taxes, (4) autonomous net exports, (5) autonomous consumption expenditure, (6) autonomous investment, and (7) financial frictions. (The use of the term *autonomous* in the factors listed above sometimes confuses students, and so it is discussed in the FYI box "What Does *Autonomous* Mean?") As we examine each case, we ask what happens to the aggregate demand curve when each of these factors changes while the inflation rate is held constant.

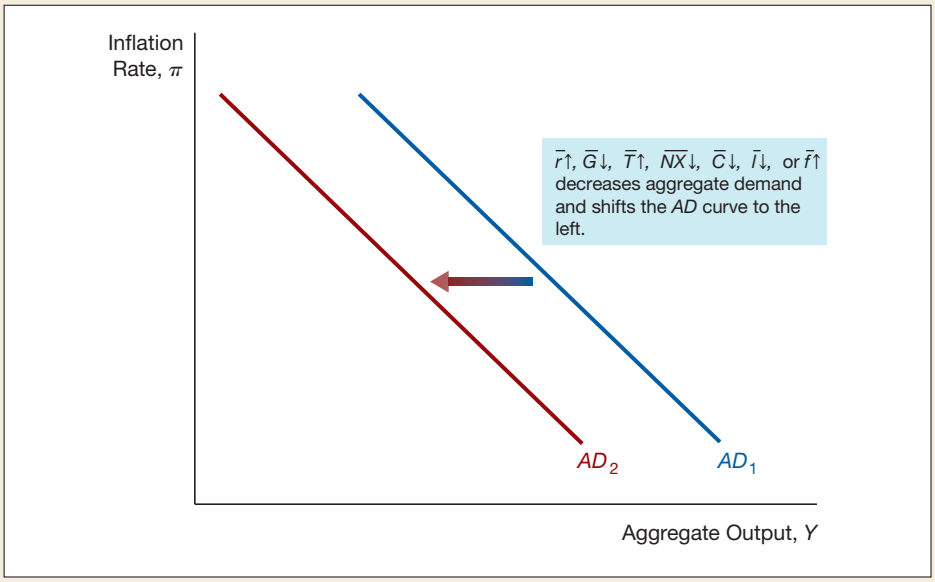
[3]If you have already read Chapters 21 and Chapter 22, this discussion of aggregate demand is just a recap of the analysis given in those chapters. Note that an additional mechanism for a downward-sloping aggregate demand curve operates through net exports, as discussed in Chapters 21 and 22.

Mini-lecture

FIGURE 23-4

Leftward Shift in the Aggregate Demand Curve

The aggregate demand curve shifts to the left from AD_1 to AD_2 when there is an autonomous tightening of monetary policy ($\bar{r}\uparrow$), a decrease in government purchases ($\bar{G}\downarrow$), an increase in taxes ($\bar{T}\uparrow$), a decrease in autonomous net exports ($\overline{NX}\downarrow$), a decrease in autonomous consumption expenditure ($\bar{C}\downarrow$), a decrease in autonomous investment ($\bar{I}\downarrow$), or an increase in financial frictions ($\bar{f}\uparrow$).



1. *Autonomous monetary policy.* We have already noted that when inflation rises, the real interest rate rises. However, the central bank often wants to make changes in the real interest rate that are autonomous, denoted by $\bar{r}$, which are movements unrelated to the variables in the model, such as the current level of the inflation rate. When the Bank of Canada decides to increase this autonomous component of the real interest rate, $\bar{r}$, the higher real interest rate at any given inflation rate leads to a higher cost of financing investment projects, which leads to a decline in investment spending and the quantity of aggregate demand, as demonstrated by the following schematic:

$$\bar{r}\uparrow \Rightarrow I\downarrow \Rightarrow Y^{ad}\downarrow$$

Therefore, aggregate demand falls at any given inflation rate, and the aggregate demand curve shifts to the left, as in Figure 23-4.

FYI What Does *Autonomous* Mean?

When economists use the word *autonomous*, they mean that this component of the variable is exogenous (independent of other variables in the model). For example, autonomous monetary policy is the component of the real interest rate set by the central bank that is unrelated to inflation or to any other variable in the model. Changes in autonomous components therefore are never associated with *movements along* a curve but rather are associated with *shifts in* the curve. Hence a change in autonomous monetary policy shifts the *AD* curve but is never a movement along the curve.

2. *Government purchases.* An increase in government purchases at any given inflation rate adds directly to aggregate demand expenditure, and hence aggregate demand rises:

$$\bar{G}\uparrow \Rightarrow Y^{ad}\uparrow$$

Aggregate demand, therefore, rises at any given inflation rate, and the aggregate demand curve shifts to the right, as in Figure 23-5.

3. *Taxes.* At any given inflation rate, an increase in taxes lowers disposable income, which leads to lower consumption expenditure and aggregate demand so that aggregate demand falls:

$$\bar{T}\uparrow \Rightarrow C\downarrow \Rightarrow Y^{ad}\downarrow$$

Aggregate demand falls at any given inflation rate, and the aggregate demand curve shifts to the left, as in Figure 23-4.

4. *Autonomous net exports.* An autonomous increase in net exports at any given inflation rate adds directly to aggregate demand and so raises aggregate demand:

$$\overline{NX}\uparrow \Rightarrow Y^{ad}\uparrow$$

Aggregate demand rises at any given inflation rate, and the aggregate demand curve shifts to the right, as in Figure 23-5.

5. *Autonomous consumption expenditure.* When consumers become more optimistic, autonomous consumption expenditure rises, and so consumers spend more at any given inflation rate. Aggregate demand therefore rises:

$$\bar{C}\uparrow \Rightarrow Y^{ad}\uparrow$$

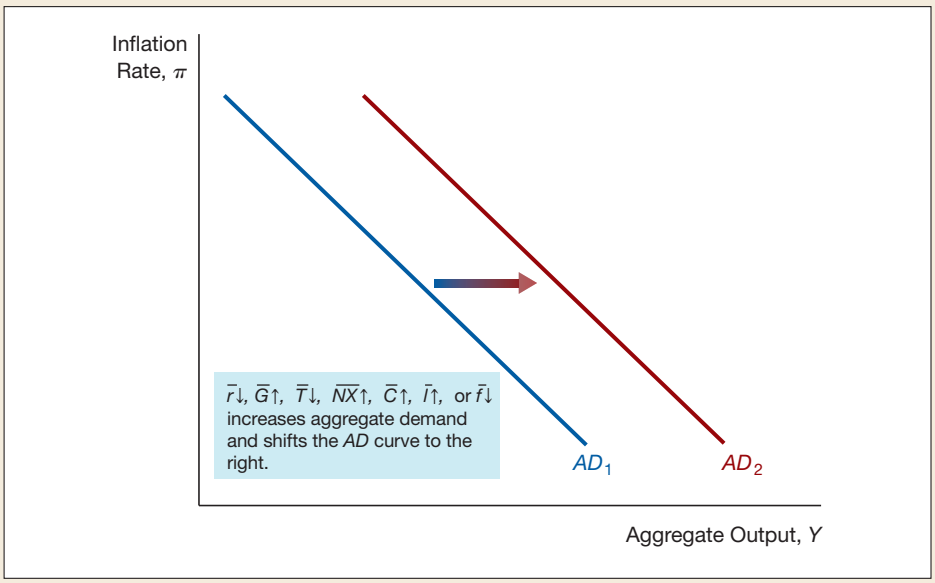
Aggregate demand rises at any given inflation rate, and the aggregate demand curve shifts to the right, as in Figure 23-5.

Mini-lecture

FIGURE 23-5

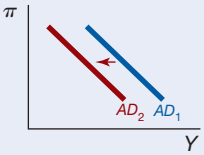
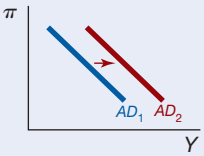
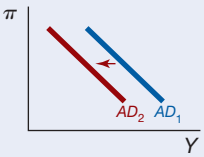
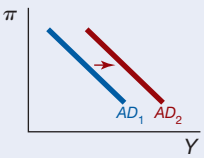
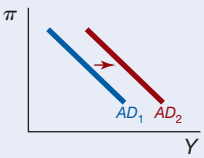
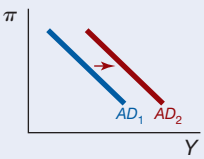
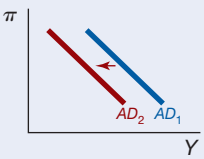
Rightward Shift in the Aggregate Demand Curve

The aggregate demand curve shifts to the right, from AD_1 to AD_2, when there is an autonomous easing of monetary policy ($\bar{r}\downarrow$), an increase in government purchases ($\bar{G}\uparrow$), a decrease in taxes ($\bar{T}\downarrow$), an increase in autonomous net exports ($\overline{NX}\uparrow$), an increase in autonomous consumption expenditure ($\bar{C}\uparrow$), an increase in autonomous investment ($\bar{I}\uparrow$), or a decrease in financial frictions ($\bar{f}\downarrow$).



SUMMARY TABLE 23-1

Factors That Shift the Aggregate Demand Curve

Factor	Change	Shift in Aggregate Demand Curve
Autonomous monetary policy, $\bar{r}$	↑	
Government purchases, $\overline{G}$	↑	
Taxes, $\overline{T}$	↑	
Autonomous net exports, $\overline{NX}$	↑	
Autonomous consumption expenditure, $\overline{C}$	↑	
Autonomous investment, $\bar{I}$	↑	
Financial frictions, $\bar{f}$	↑	

Note: Only increases (↑) in the factors are shown. The effect of decreases in the factors would be the opposite of those indicated in the "Shift" column.

6. *Autonomous investment.* When businesses become more optimistic, autonomous investment rises and businesses spend more at any given inflation rate. Planned investment increases and aggregate demand rises:

$$\bar{I}\uparrow \Rightarrow Y^{ad}\uparrow$$

Aggregate demand rises at any given inflation rate, and the aggregate demand curve shifts to the right, as in Figure 23-5.

7. *Financial frictions.* The real cost of borrowing reflects not only the real interest rate on default-free debt instruments, r, but also financial frictions, denoted by $\bar{f}$, which are additions to the real cost of borrowing caused by asymmetric information problems in financial markets (described in Chapter 8). When financial frictions increase, the real cost of borrowing increases, so planned investment spending falls at any given inflation rate, and aggregate demand falls:

$$\bar{f}\uparrow \Rightarrow I\downarrow \Rightarrow Y^{ad}\downarrow$$

Aggregate demand falls at any given inflation rate, and the aggregate demand curve shifts to the left as in Figure 23-4.

The conclusion from our analysis is as follows: ***Aggregate demand increases at any given inflation rate, and the aggregate demand curve shifts to the right when there is (1) an autonomous easing of monetary policy*** ($\bar{r}\downarrow$), ***(2) an increase in government purchases*** ($\overline{G}\uparrow$), ***(3) a decrease in taxes*** ($\overline{T}\downarrow$), ***(4) an increase in autonomous net exports*** ($\overline{NX}\uparrow$), ***(5) an increase in autonomous consumption expenditure*** ($\overline{C}\uparrow$), ***(6) an increase in autonomous investment*** ($\bar{I}\uparrow$), ***or (7) a decrease in financial frictions*** ($\bar{f}\downarrow$). ***Conversely, the aggregate demand curve shifts to the left when any of these factors change in the opposite direction.*** As a study aid, Summary Table 23-1 summarizes the shifts in the aggregate demand curve that occur with changes in each of these seven factors.

23.3 AGGREGATE SUPPLY

LO 23.3 Illustrate and interpret the short-run and long-run aggregate supply curves.

To complete our analysis, we need to derive an **aggregate supply curve**, a curve that shows the relationship between the quantity of output supplied and the inflation rate. In the typical supply and demand analysis, we have only one supply curve, but because prices and wages take time to adjust to their long-run levels, the aggregate supply curve differs in the short and long runs. First we examine the long-run aggregate supply curve; then we derive the short-run aggregate supply curve. Finally, we look at how both of these curves shift over time and at how the economy moves from the short run to the long run.

Long-Run Aggregate Supply Curve

The amount of output that can be produced in the economy in the long run is determined by the amount of capital in the economy, the amount of labour supplied at full employment, and the available technology. As discussed in Chapter 17, some unemployment cannot be helped because it is either frictional or structural. Thus, at full employment, unemployment is not at zero but is rather at a level above zero at which the labour market is in equilibrium. This **natural rate of unemployment** is the rate to

Mini-lecture

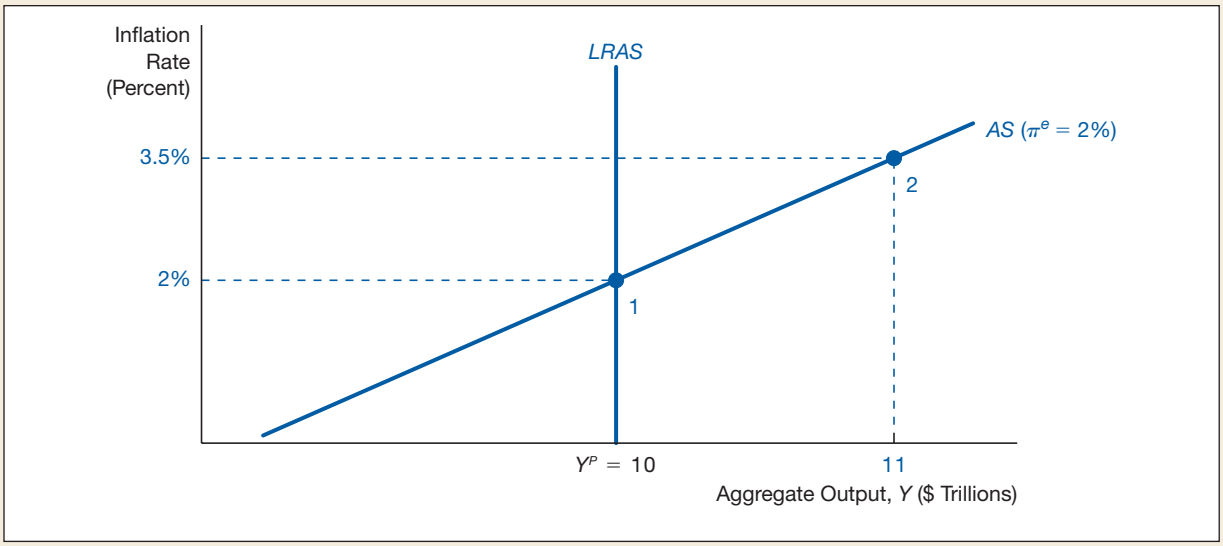


FIGURE 23-6 Long- and Short-Run Aggregate Supply Curves

The amount of aggregate output supplied at any given inflation rate is at potential output in the long run, $10 trillion, so that the long-run aggregate supply curve *LRAS* is a vertical line at Y^P = $10 trillion. The short-run aggregate supply curve, *AS*, where π^e = 2%, is upward-sloping because as *Y* rises relative to Y^P, labour markets get tighter and inflation rises. *AS* intersects *LRAS* at point 1, where current inflation equals expected inflation of 2%.

which the economy gravitates in the long run. Many economists believe that the natural rate of unemployment is currently around 6%.

The level of aggregate output produced at the natural rate of unemployment is called the **natural rate of output** but is more often referred to as **potential output**: It is the level of output generated when all firms on average are operating at their long-run, profit-maximizing level, and it is where the economy settles in the long run for any inflation rate. Suppose potential output is at $10 trillion. The long-run aggregate supply curve (*LRAS*) is then vertical at the $10 trillion of potential output, denoted by Y^P, as drawn in Figure 23-6.

Short-Run Aggregate Supply Curve

The short-run aggregate supply curve is based on the idea that three factors drive inflation: (1) expectations of inflation, (2) the output gap, and (3) inflation (supply) shocks.

Expected Inflation, π^e Workers and firms care about wages in real terms—that is, in terms of the goods and services that wages can buy. When workers expect a positive inflation rate, they will adjust nominal wages upward one-to-one with the expected inflation rate so that the real wage rate does not decrease. Thus, holding everything else constant, wage inflation will rise one-to-one with rises in expected inflation. Because

wages are the most important cost of producing goods and services, overall inflation will also rise on a one-to-one basis with increases in expected inflation.

Output Gap The **output gap** is defined as the percentage difference between aggregate output and potential output, $Y - Y^P$. When output exceeds its potential level and the output gap is positive, there is very little slack in the economy. Workers will demand higher wages, and firms will take the opportunity to raise prices. The end result will be higher inflation. Conversely, when the output gap is negative, there will be a lot of slack in the economy. Thus workers will accept smaller increases in wages, and firms will need to lower prices to sell their goods, resulting in lower inflation.

Inflation (Supply) Shocks **Supply shocks** occur when there are shocks to the supply of goods and services produced in the economy that translate into **inflation shocks**, that is, shifts in inflation that are independent of the amount of slack in the economy or expected inflation. For example, when the supply of oil is restricted, as has occurred several times when Middle East countries were at war, the resulting rise in the price of oil leads firms to raise prices more to reflect the increased costs of production, thus driving up inflation. Energy inflation shocks can also occur when demand increases—for example, increased demand from developing countries like China such as occurred in 2007–2008—again driving up inflation. Inflation shocks can also come from rises in import prices or from **cost-push shocks** in which workers push for wages that are higher than productivity gains, thereby driving up costs and inflation.

Short-Run Aggregate Supply Curve Putting all of our analysis together, we can write the following equation for the short-run aggregate supply curve:

$$\begin{array}{lcl} \pi & = & \pi^e + \gamma(Y - Y^P) + \rho \\ \text{inflation} & = & \text{expected inflation} + \gamma \times \text{output gap} + \text{inflation shock} \end{array} \qquad (2)$$

where
π = inflation
π^e = expected inflation
$Y - Y^P$ = the output gap
γ = the sensitivity of inflation to the output gap
ρ = the inflation shock

The short-run aggregate supply curve given in Equation 2 tells us that inflation is driven by three factors: (1) expectations of inflation, (2) output gaps, and (3) inflation shocks.

Why the Short-Run Aggregate Supply Curve Is Upward-Sloping To see why the short-run aggregate supply curve, *AS*, in Figure 23-6 is upward-sloping, let's assume that expected inflation is at 2% and that there are no inflation shocks. When actual output equals potential output at $10 trillion, the output gap, $Y - Y^P$, is zero, and so Equation 2 indicates that the inflation rate will equal the expected inflation rate of 2%. The combination of $10 trillion of aggregate output and a 2% inflation rate is shown by point 1 on the *AS* curve. [Note that the short-run aggregate supply curve in Figure 23-6 is marked as *AS* $(\pi^e = 2\%)$ to indicate it is drawn assuming $\pi^e = 2\%$.]

Now suppose that aggregate output rises to \$11 trillion. Because there is a positive output gap ($Y = \$11$ trillion $> Y^P = \$10$ trillion), Equation 2 indicates that inflation will rise above 2%, say, to 3.5%, marked as point 2. The curve connecting points 1 and 2 is the short-run aggregate supply curve, *AS*, and it is upward-sloping. When Y rises relative to Y^P and $Y > Y^P$, the labour market is tighter and firms raise their prices at a more rapid rate, causing the inflation rate to rise. Thus the *AS* curve is upward-sloping.

Price Stickiness and the Short-Run Aggregate Supply Curve

In the short run, wages and prices are **sticky**, which means that wages and prices are slow to adjust. The more flexible wages and prices are—that is, the less sticky they are—the more rapidly they, and inflation, respond to deviations of output from potential output; that is, more flexible wages and prices imply that the value of γ is higher, which in turn implies that the short-run aggregate supply curve is steeper. If wages and prices are completely flexible, then γ becomes so large that the short-run aggregate supply curve becomes vertical and is identical to the long-run aggregate supply curve.

23.4 SHIFTS IN THE AGGREGATE SUPPLY CURVES

LO 23.4 Illustrate and interpret shifts in the short-run and long-run aggregate supply curves.

Now that we have examined the long-run and short-run aggregate supply curves, we can look at why each of these curves shifts.

Shifts in the Long-Run Aggregate Supply Curve

The quantity of output supplied in the long run is determined by the three factors that cause potential output to change and thus shift the long-run aggregate supply curve: (1) the total amount of capital in the economy, (2) the total amount of labour supplied in the economy, and (3) the available technology that puts labour and capital together to produce goods and services. When any of these three factors increases, potential output rises, say, from $Y_1^P = \$10$ trillion to $Y_2^P = \$11$ trillion, and the long-run aggregate supply curve shifts to the right, from $LRAS_1$ to $LRAS_2$ in Figure 23-7.

Because all three of these factors typically grow fairly steadily over time, Y^P and the long-run aggregate supply curve will keep shifting to the right at a steady pace. To keep things simple in later diagrams in this and following chapters, when Y^P is growing at a steady rate, we will represent Y^P and the long-run aggregate supply curve as fixed.

Another source of shifts in the long-run aggregate supply curve is changes in the natural rate of unemployment. If the natural rate of unemployment declines, it means that labour is being more heavily utilized, and so potential output will increase. A decline in the natural rate of unemployment thus shifts the long-run aggregate supply curve to the right, from $LRAS_1$ to $LRAS_2$ in Figure 23-7. A rise in the natural rate of unemployment would have the opposite effect, shifting the long-run aggregate supply curve to the left.

The conclusion from our analysis is as follows: ***The long-run aggregate supply curve shifts to the right when there is (1) an increase in the total amount of capital in the economy, (2) an increase in the total amount of labour supplied in the economy, (3) an increase in available technology, or (4) a decline in the natural rate of unemployment. An opposite movement in these variables shifts the LRAS curve to the left.***

Mini-lecture

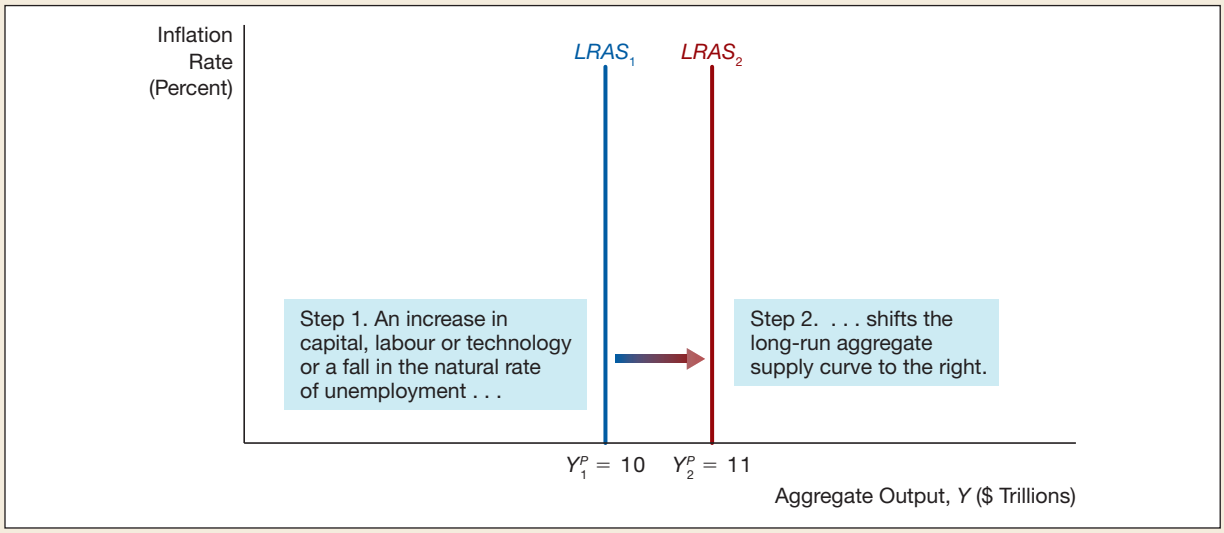

Inflation Rate (Percent)
LRAS1
LRAS2
Step 1. An increase in capital, labour or technology or a fall in the natural rate of unemployment . . .
Step 2. . . . shifts the long-run aggregate supply curve to the right.
Y1P = 10
Y2P = 11
Aggregate Output, Y ($ Trillions)

FIGURE 23-7 Shift in the Long-Run Aggregate Supply Curve

The long-run aggregate supply curve shifts to the right from $LRAS_1$ to $LRAS_2$ when there is (1) an increase in the total amount of capital in the economy, (2) an increase in the total amount of labour supplied in the economy, (3) an increase in the available technology, or (4) a decline in the natural rate of unemployment. An opposite movement in these variables shifts the *LRAS* curve to the left.

Shifts in the Short-Run Aggregate Supply Curve

The three terms on the right-hand side of Equation 2 suggest that three factors can shift the short-run aggregate supply curve: (1) expected inflation, (2) a persistent output gap, and (3) inflation shocks.

Expected Inflation What if a newly appointed governor of the Bank of Canada does not think that inflation is costly and so is willing to tolerate an inflation rate that is two percentage points higher than the current rate? Households and firms will then expect that the Bank of Canada will pursue policies that will let inflation rise by, say, two percentage points in the future and will want to raise wages and prices by this additional amount. In such a situation, expected inflation π^e will jump by two percentage points, and the short-run aggregate supply curve will shift upward and to the left, from AS_1 to AS_2 in Figure 23-8. ***A rise in expected inflation causes the short-run aggregate supply curve to shift upward and to the left. Conversely, a fall in expected inflation causes the short-run aggregate supply curve to shift down and to the right. The larger the change in expected inflation, the larger is the shift.***

Inflation Shock Suppose that energy prices suddenly shoot up because terrorists destroy a number of oil fields. This supply restriction (referred to as a *negative supply shock*) causes the inflation-shock term ρ in Equation 2 to increase, and so the short-run aggregate supply curve shifts up and to the left, from AS_1 to AS_2 in Figure 23-8. In contrast, a favourable (positive) supply shock drives down inflation and so has the opposite

Mini-lecture

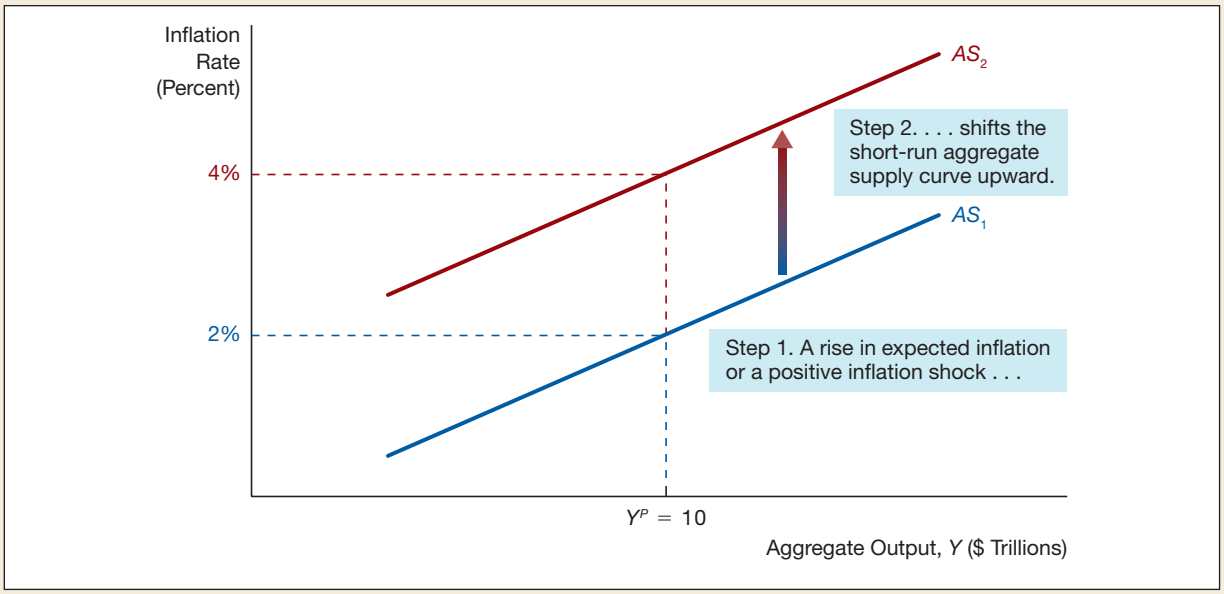

Inflation Rate (Percent)
4%
2%
AS_2
AS_1
Step 2. . . . shifts the short-run aggregate supply curve upward.
Step 1. A rise in expected inflation or a positive inflation shock . . .
$Y^P = 10$
Aggregate Output, Y ($ Trillions)

FIGURE 23-8 Shift in the Short-Run Aggregate Supply Curve from Changes in Expected Inflation and Supply (Inflation) Shocks

A rise in expected inflation or a negative supply shock (a positive inflation shock) shifts the short-run aggregate supply curve upward from AS_1 to AS_2. (A decrease in expected inflation or a negative inflation shock would lead to a downward shift of the *AS* curve.)

effect: It causes the short-run aggregate supply curve to shift down and to the right. ***Negative supply shocks drive up inflation and cause the short-run aggregate supply curve to shift up and to the left, while positive supply shocks lower inflation and cause the short-run aggregate supply curve to shift down and to the right.***

Persistent Output Gap We have already seen that a higher output gap leads to higher inflation, causing a movement along the short-run aggregate supply curve. We can represent this scenario by the movement from point 1 to point 2 on the initial short-run aggregate supply curve AS_1 in Figure 23-9. A persistent output gap, however, will cause the short-run aggregate supply curve to shift by affecting expected inflation. To see this, consider what happens if the aggregate output stays at $11 trillion, which is greater than the potential output of $Y^P = \$10$ trillion, so that the output gap remains persistently positive. At point 2 on the initial short-run aggregate supply curve, AS_1, output has risen to $11 trillion and inflation has risen from 2% to 3.5%. This higher level of inflation will cause expected inflation to rise during the next period, and so the short-run aggregate supply curve for the next period, AS_2, will shift upward. If output remains above potential output at point 3, inflation will rise further, to 5.0%. This higher inflation will then lead to higher expected inflation and, as the vertical arrow indicates, the short-run aggregate supply curve will shift upward to AS_3 during the next period.

When will the short-run aggregate supply curve stop rising? It will stop only when output returns to its potential level and the output gap disappears. At this point, there

Mini-lecture

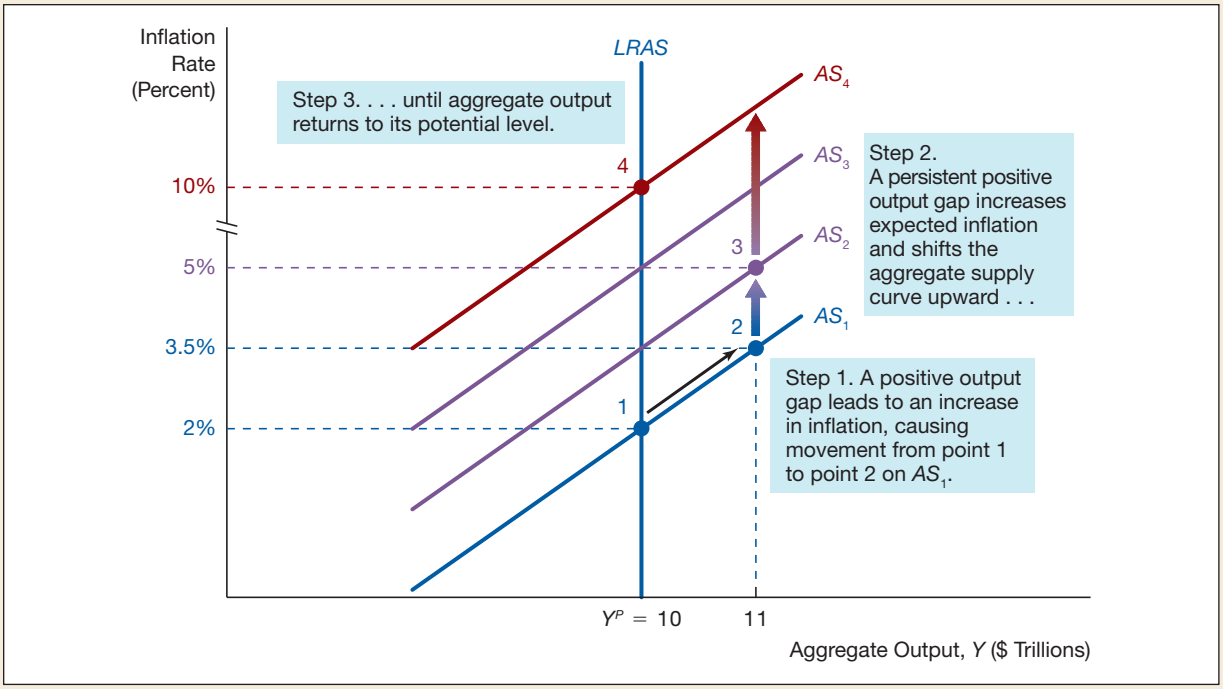

Inflation Rate (Percent)
LRAS
AS_4
AS_3
AS_2
AS_1
Step 3. . . . until aggregate output returns to its potential level.
Step 2. A persistent positive output gap increases expected inflation and shifts the aggregate supply curve upward . . .
Step 1. A positive output gap leads to an increase in inflation, causing movement from point 1 to point 2 on AS_1.
10%
5%
3.5%
2%
4
3
2
1
Y^P = 10
11
Aggregate Output, Y ($ Trillions)

FIGURE 23-9 Shift in the Short-Run Aggregate Supply Curve from a Persistent Positive Output Gap

When output is above potential, the economy moves along the AS_1 curve from point 1 to point 2 and inflation rises to 3.5%. If output continues to remain above potential output, where the output gap is positive, the short-run aggregate supply curve will shift upward, to AS_2 and then to AS_3. The short-run aggregate supply curve stops shifting upward when the economy reaches point 4 on the short-run aggregate supply curve AS_4, where the output gap is once again zero.

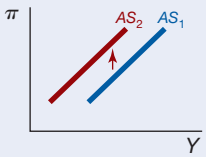
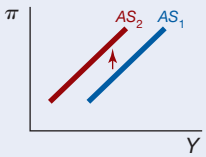
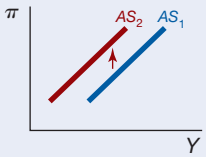
is no longer a reason for actual and hence expected inflation to rise. Suppose this happens when inflation is at 10% and aggregate output Y = \$10 trillion = Y^P. Because the output gap is now zero, the aggregate supply curve drawn through point 4, AS_4, has no reason to shift because inflation and expected inflation have stopped rising.

The same reasoning indicates that if aggregate output is kept below potential for a period of time, $Y < Y^P$, then the short-run aggregate supply curve will shift downward and to the right. This downward shift of the aggregate supply curve will stop only when output returns to its potential level and the economy is back on the long-run aggregate supply curve.

Our analysis yields the following conclusion: ***When aggregate output is above potential output, so that a persistent positive output gap exists, the short-run aggregate supply curve shifts up and to the left. Conversely, when aggregate output falls below potential output, the short-run aggregate supply curve shifts down and to the right. Only when aggregate output returns to potential output does the short-run aggregate supply curve stop shifting.*** As a study aid, Summary Table 23-2 summarizes the shifts in the short-run aggregate supply curve caused by each of these three factors.

SUMMARY TABLE 23-2

Factors That Shift the Short-Run Aggregate Supply Curve

Factor	Change	Shift in Supply Curve
Expected inflation, π^e	↑	
Inflation shock, ρ	↑	
Persistent output gap, $(Y - Y^P)$	↑	

Note: Only increases (↑) in the factors are shown. The effect of decreases in the factors would be the opposite of those indicated in the "Shift" column.

23.5 EQUILIBRIUM IN AGGREGATE DEMAND AND SUPPLY ANALYSIS

LO 23.5 Illustrate and interpret the short-run and long-run equilibria and the role of the self-correcting mechanism.

We can now put the aggregate demand and supply curves together to describe **macroeconomic equilibrium** in the economy, the point at which all markets are simultaneously in equilibrium and the quantity of aggregate output demanded equals the quantity of aggregate output supplied. We represent macroeconomic equilibrium graphically as the point at which the aggregate demand curve intersects the aggregate supply curve. However, recall that we have two aggregate supply curves: one for the short run and one for the long run. Consequently, in the context of aggregate supply and demand analysis, there are short-run and long-run equilibria. In this section, we illustrate equilibrium in the short and long runs. In following sections, we examine aggregate demand and aggregate supply shocks that lead to changes in equilibrium.[4]

[4]If you are more mathematically inclined and would like to see how to solve for the macroeconomic equilibrium using algebra, go to the first online appendix to Chapter 23.

Mini-lecture

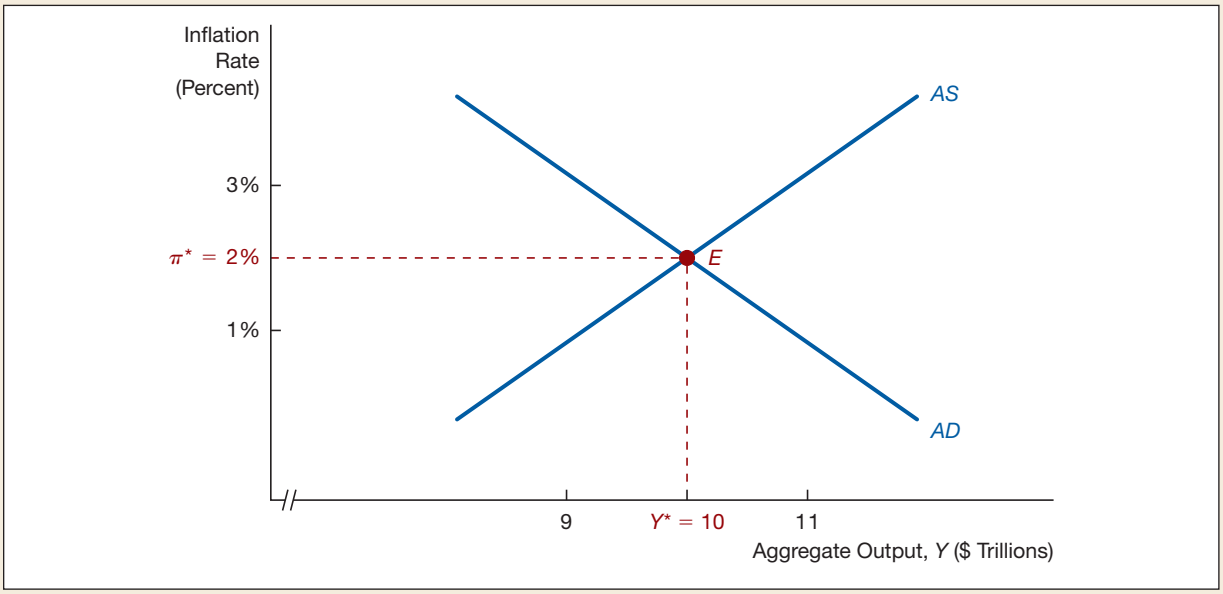


FIGURE 23-10 Short-Run Equilibrium

Short-run equilibrium occurs at point *E*, at the intersection of the aggregate demand curve *AD* and the short-run aggregate supply curve *AS*.

Short-Run Equilibrium

Figure 23-10 is an ***AD/AS* diagram** that shows the aggregate demand curve (*AD*), the short-run aggregate supply curve (*AS*), and the macroeconomic equilibrium. It illustrates a short-run equilibrium in which the quantity of aggregate output demanded equals the quantity of output supplied. In Figure 23-10, the short-run aggregate demand curve *AD* and the short-run aggregate supply curve *AS* intersect at point *E*, with an equilibrium level of aggregate output at $Y^* = \$10$ trillion and an equilibrium inflation rate at $\pi^* = 2\%$.

Aggregate Demand and Supply Analysis Using an Aggregate Output Index

As we demonstrated in Figure 23-10, potential output grows steadily over time, and so the long-run aggregate supply curve keeps shifting to the right over time. So, knowing an economy's equilibrium level of aggregate output is useful, but it does not tell us if the economy is doing as well as it could. For example, in the third quarter of 2007, just before the global financial crisis started, Canadian real GDP was $1.73 trillion in 2012 dollars. This number grew to $1.84 trillion in the second quarter of 2020. However, knowing that real GDP grew over this thirteen-year period does not tell us whether the economy was living up to its potential.

As we saw earlier in the chapter, economists are particularly interested in business cycle fluctuations, which are short-run deviations of aggregate output around potential output, which is measured by the output gap, the percentage difference between actual real GDP and potential GDP.

In the third quarter of 2007, just before the recession began, the output gap was +1.8%, but it fell to −10.2% by the second quarter of 2020, even though real GDP had risen over the same period by 6.3%, from $1.73 trillion to $1.84 trillion in 2012 dollars. To examine business cycle fluctuations, it is useful to modify the *AD/AS* diagram to show the level of equilibrium aggregate output relative to potential GDP (and therefore the output gap). To modify the *AD/AS* diagram, we express real GDP in terms of an **aggregate output index**, in which potential GDP is designated as 100 and all other values of real GDP are set relative to 100.

Calculating an Aggregate Output Index. An aggregate output index is calculated as

$$\text{Aggregate Output Index} = \frac{\text{Aggregate Output}}{\text{Potential Output}} \times 100$$

To see how the aggregate output index is calculated, suppose that potential GDP is $10 trillion, then

- At $9 trillion, the aggregate output index is 90 (=[$9 trillion/$10 trillion] × 100 = 0.90 × 100). In other words, aggregate output is 10% below potential output and the output gap is −10%.
- At $10 trillion, which equals potential GDP, the aggregate output index is 100 (=[$10 trillion/$10 trillion] × 100 = 1.00 × 100). In other words, aggregate output equals potential output and the output gap is zero.
- At $11 trillion, the aggregate output index is 110 (=[$11 trillion/$10 trillion] × 100 = 1.10 × 100). In other words, aggregate output is 10% above potential output and the output gap is +10%.

An *AD/AS* Diagram with an Aggregate Output Index In Figure 23-11, we draw an *AD/AS* diagram with an aggregate output index. We replace aggregate output on the horizontal axis with the aggregate output index.

We replace the aggregate output value of $9 trillion with its aggregate output index value of 90. The aggregate output value of $10 trillion, which equals potential GDP, is replaced by its aggregate output index value of 100. Aggregate output of $11 trillion is replaced by its aggregate output index value of 110. The equilibrium at point *E* now indicates that the equilibrium aggregate output $Y^* = 100$, where, as in the previous figure, aggregate output is at the $10 trillion, potential output level.

The value of the output gap is readily obtained from the aggregate output index, as is shown by the boxes in the figure. At the equilibrium, the aggregate output index $Y^* = 100$, so the output gap is zero. When the value of the aggregate output index is 90, aggregate output is 10% below potential GDP, so the output gap is −10%. When the value of the aggregate output index is 110, aggregate output is 10% above potential output, so the output gap is +10%.

How the Short-Run Equilibrium Moves to the Long-Run Equilibrium over Time

In supply and demand analysis, once we find the equilibrium point at which the quantity demanded equals the quantity supplied, additional analysis typically is not needed. In *aggregate* supply and demand analysis, however, that is not the case. Even when the

Mini-lecture

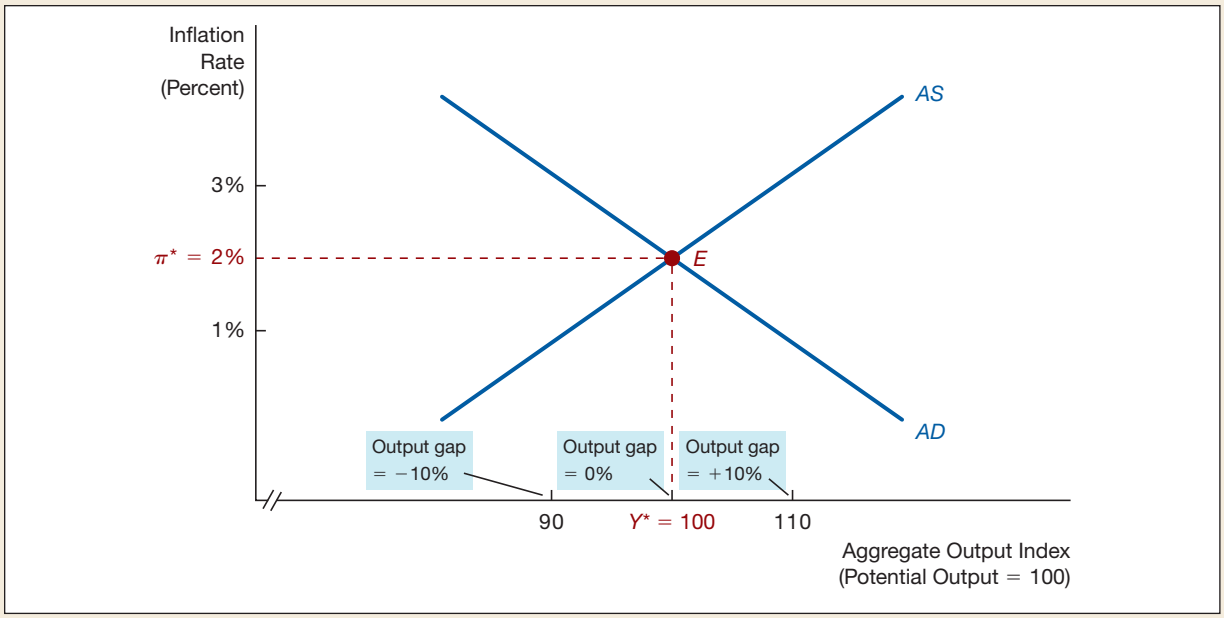


FIGURE 23-11 ***AD/AS* Diagram Using an Aggregate Output Index**

With the horizontal axis changed to an aggregate output index and potential output of $10 trillion, the aggregate output value of $9 trillion is replaced by an aggregate output index of 90, the aggregate output value of $10 trillion is replaced by the aggregate output index of 100, and the aggregate output value of $11 trillion is replaced by the aggregate output index of 110. The corresponding output gap that corresponds to each value of the aggregate output index is also shown.

quantity of aggregate output demanded equals the quantity supplied at the intersection of the aggregate demand curve and the short-run aggregate supply curve, if output differs from its potential level ($Y^* \neq Y^P$), the short-run equilibrium will move over time. To understand why, recall that if the current level of inflation changes from its initial level, the short-run aggregate supply curve will shift as wages and prices adjust to a new expected rate of inflation.

We look at how the short-run equilibrium changes over time in response to two situations: when short-run equilibrium output is initially above potential output (the natural rate of output) and when it is initially below potential output.

In panel (a) of Figure 23-12, the initial equilibrium occurs at point 1, the intersection of the aggregate demand curve *AD* and the initial short-run aggregate supply curve AS_1. The long-run aggregate supply curve, *LRAS*, is a vertical line at an aggregate output index of 100, because on the long-run aggregate supply curve, aggregate output is always equal to potential output at any given inflation rate. The level of the aggregate output index at point 1 is 110, so aggregate output is greater than potential output, and there is excessive tightness in the labour market. Hence, the positive output gap at 110 drives wages up and causes firms to raise their prices at a more rapid rate. Inflation then rises above the initial inflation rate, say, to 2.5%. At this higher inflation rate, firms and households adjust their expectations next period, and expected inflation becomes higher. Wages and prices then rise more rapidly, and the aggregate supply curve shifts up and to the left, from AS_1 to AS_2.

Mini-lecture

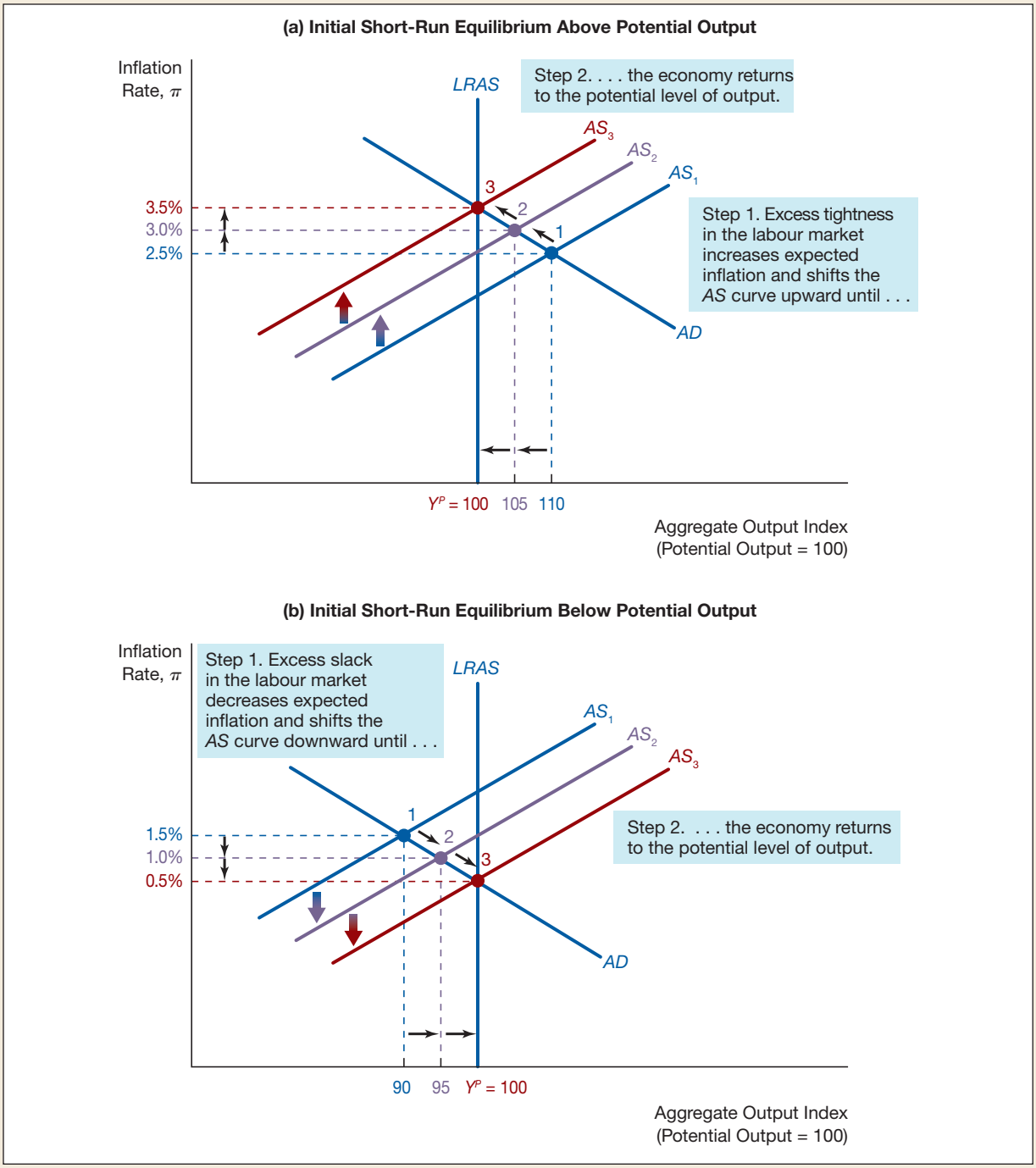


FIGURE 23-12 Adjustment to Long-Run Equilibrium in Aggregate Supply and Demand Analysis

In both panels, the initial short-run equilibrium is at point 1, at the intersection of *AD* and AS_1. In panel (a), initial short-run equilibrium is above potential output, the long-run equilibrium, and so the short-run aggregate supply curve shifts upward until it reaches AS_3, where the aggregate output index returns to 100 and inflation rises to 3.5%. In panel (b), the initial short-run equilibrium is below potential output, and so the short-run aggregate supply curve shifts downward until the aggregate output index returns to 100 and inflation falls to 0.5%. In both panels, the economy's self-correcting mechanism returns the economy to the level of potential output.

The new short-run equilibrium at point 2 is an upward movement along the aggregate demand curve, and the aggregate output index falls to 105 and inflation rises to 3.0%. However, because aggregate output is still above potential output, inflation continues to be above its value last period. Expected inflation rises further, eventually shifting the aggregate supply curve up and to the left to AS_3. The economy reaches long-run equilibrium at point 3 on the vertical long-run aggregate supply curve (*LRAS*), at an aggregate output index of 100 and an inflation rate of 3.5%. Because output now is at potential, there is no further pressure on inflation to rise and thus no further tendency for the aggregate supply curve to shift.

The movements in panel (a) indicate that the economy will not remain at a level of output higher than potential output over time. Specifically, the short-run aggregate supply curve will shift up and to the left, raising the inflation rate and causing the economy (equilibrium) to move upward along the aggregate demand curve until it comes to rest at a point on the long-run aggregate supply curve (*LRAS*) that is at potential output, where the aggregate output index is 100.

In panel (b), at the initial equilibrium at point 1, the aggregate output index is at 90, which is 10% below the level of potential output. Because unemployment is now above its natural rate, there is excess slack in the labour markets. This slack at an aggregate output index of 90 decreases inflation to 1.5%, which decreases expected inflation and shifts the short-run aggregate supply curve in the next period down and to the right, to AS_2.

The equilibrium now moves to point 2, and the aggregate output index rises to 95, while the inflation rate falls to 1.0%. However, because aggregate output at an aggregate output index of 95 is still below potential output, inflation again is below its value last period, causing expected inflation to fall further. This decline in expected inflation shifts the aggregate supply curve down until it comes to rest at AS_3. The economy (equilibrium) moves downward along the aggregate demand curve until it reaches the long-run equilibrium at point 3, at the intersection of the aggregate demand curve (*AD*) and the long-run aggregate supply curve (*LRAS*) at an aggregate output index of 100 and an inflation rate of 0.5%. Here, as in panel (a), the economy comes to rest when output has again returned to its potential level.

Self-Correcting Mechanism

Notice that in both panels of Figure 23-12, regardless of where output is initially, it returns eventually to potential output, a feature we call the **self-correcting mechanism**. The self-correcting mechanism occurs because the short-run aggregate supply curve shifts up or down to restore the economy to the long-run equilibrium at full employment (aggregate output at potential) over time.

23.6 CHANGES IN EQUILIBRIUM: AGGREGATE DEMAND SHOCKS

LO 23.6 Illustrate and interpret the short-run and long-run effects of a shock to aggregate demand.

With an understanding of the distinction between the short-run and long-run equilibria, we are now ready to analyze what happens when the economy experiences demand shocks, shocks that cause the aggregate demand curve to shift. Figure 23-13 depicts the effects of a rightward shift in the aggregate demand curve due to positive demand shocks, which may be caused by the following:

- An autonomous easing of monetary policy ($\bar{r}\downarrow$, a lowering of the real interest rate at any given inflation rate)
- An increase in government purchases ($\bar{G}\uparrow$)
- A decrease in taxes ($\bar{T}\downarrow$)
- An increase in autonomous net exports ($\overline{NX}\uparrow$)
- An increase in autonomous consumption expenditure ($\bar{C}\uparrow$)
- An increase in autonomous investment ($\bar{I}\uparrow$)
- A decrease in financial frictions ($\bar{f}\downarrow$)

Figure 23-13 shows the economy initially in long-run equilibrium at point 1, where the initial aggregate demand curve AD_1 intersects the short-run aggregate supply AS_1 curve at an aggregate output index of 100 and the inflation rate is 2%. Suppose that consumers and businesses become more optimistic and that the resulting increases in autonomous consumption and investment create a positive demand shock that shifts the aggregate demand curve to the right to AD_2. The economy moves up the short-run aggregate supply curve AS_1 to point 2, and the aggregate output index and inflation rise to 110 and 3.5%, respectively. However, the economy will not remain at point 2 in the long run because the aggregate output index at 110 is above potential output.

Mini-lecture

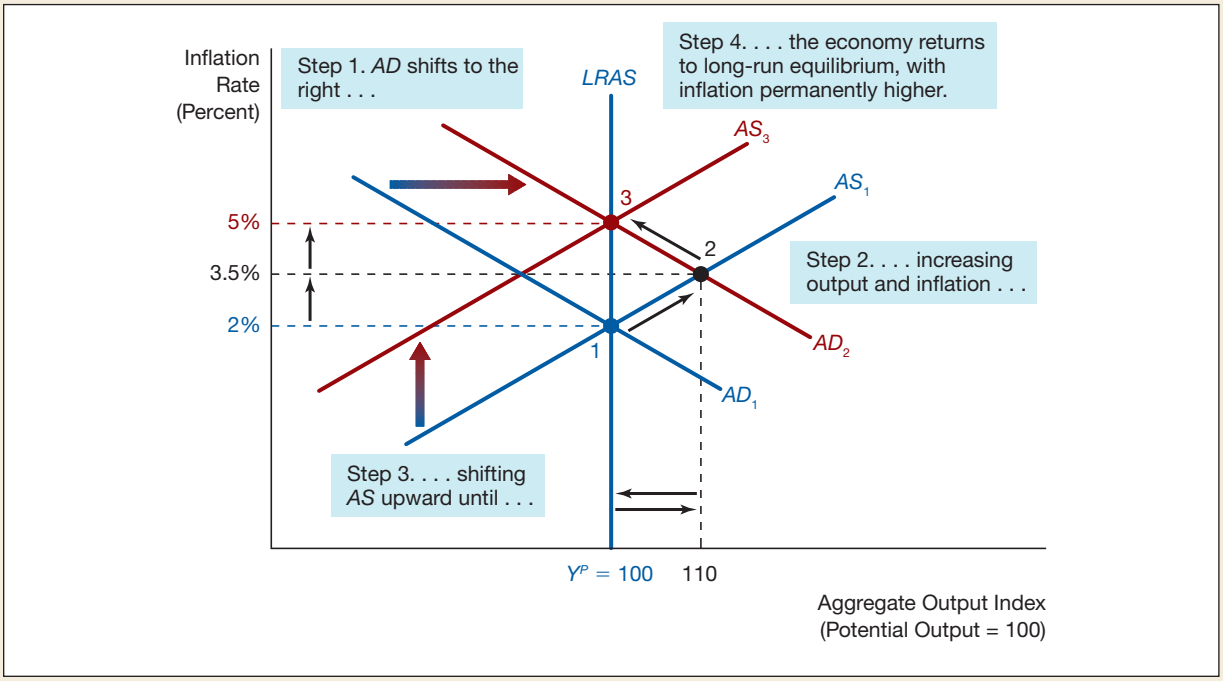


FIGURE 23-13 Positive Demand Shock

A positive demand shock shifts the aggregate demand curve upward from AD_1 to AD_2 and moves the economy from point 1 to point 2, resulting in higher inflation at 3.5% and a higher aggregate output index of 110. Because output is greater than potential output and therefore expected inflation increases, the short-run aggregate supply curve begins to shift up, eventually reaching AS_3. At point 3, the economy returns to long-run equilibrium, with the aggregate output index at 100 and the inflation level rising to 5%.

Expected inflation will rise, and the short-run aggregate supply curve will eventually shift upward to AS_3. The economy (equilibrium) thus will move up the AD_2 curve from point 2 to point 3, which is the point of long-run equilibrium at which inflation equals 5% and output returns to potential with the aggregate output index at 100. ***Although the initial short-run effect of the rightward shift in the aggregate demand curve is a rise in both inflation and output, the ultimate long-run effect is a rise in inflation only, because the aggregate output index returns to its initial level of 100.***[5]

We now apply the aggregate demand and supply model to demand shocks, as a payoff for our hard work in constructing the model. Throughout the remainder of this chapter, we will apply aggregate supply and demand analysis to a number of business cycle episodes. To simplify our analysis, we will assume in all examples that aggregate output is initially at the level of potential output.

APPLICATION The Bank of Canada Disinflation, 1981–1985

In the early 1980s, inflation in Canada had spun out of control and the inflation rate exceeded 12%. The Bank of Canada was determined to get inflation down. The inflation rate fell from 12.5% in 1981 to 5.8% in 1983, and then to 3.9% in 1985. But the decline in inflation came at a high cost: The economy experienced its worst recession since World War II, with the aggregate output index falling to 94.6 in 1983 (which corresponded to an output gap of −5.4% and an unemployment rate of 11.9%). By 1985, the economy had recovered and the aggregate output index rose to 98.7, where aggregate output was close to its potential output level.

Figure 23-14 shows how aggregate demand and supply analysis can explain these outcomes. Note that to keep this and later figures in applications less cluttered, we no longer draw the long-run aggregate supply curve in the AD/AS diagram, because it is always just a vertical line at an aggregate output index of 100. Instead, there is a tick mark at an aggregate output index of 100 to remind you where the long-run aggregate would otherwise be.

The initial equilibrium is at point *1981*. The initial equilibrium in 1981 occurred at the intersection of the aggregate demand curve, AD_{1981}, and the short-run aggregate supply curve, AS_{1981}. The resulting inflation rate in 1981 was 12.5%, and the aggregate output index was 103.8, so the economy was close to the long-run equilibrium.

The monetary tightening led to a higher real interest rate at any given inflation rate. As a result, consumption expenditure and investment spending fell precipitously, leading to a decrease in the quantity of aggregate output demanded at any given inflation rate. By 1983, the monetary tightening that started in 1981 and continued through 1983 had shifted the aggregate demand curve to the left to AD_{1983}. The equilibrium moved down the short-run aggregate supply curve to point *1983*, at the intersection of the initial short-run aggregate supply curve, AS_{1981}, and the aggregate demand curve, AD_{1983}. The equilibrium inflation rate then fell to 5.8% in 1983, while the equilibrium output index fell to 94.6, that is, 5.4% below potential output. The resulting recession was very severe, with the unemployment rate rising from an average of 7.6% in 1981 to 11.9% in 1983.

[5]The analysis here assumes that each of these positive demand shocks occurs while everything else is held constant, the usual *ceteris paribus* ("other things being equal") assumption that is standard in supply and demand analysis. Specifically, this means that the central bank is assumed not to be responding to demand shocks. In the next chapter, we relax this assumption and allow monetary policymakers to respond to these shocks. As we will see, if monetary policymakers want to keep inflation caused by a positive demand shock from rising, they will respond by autonomously tightening monetary policy.

Mini-lecture

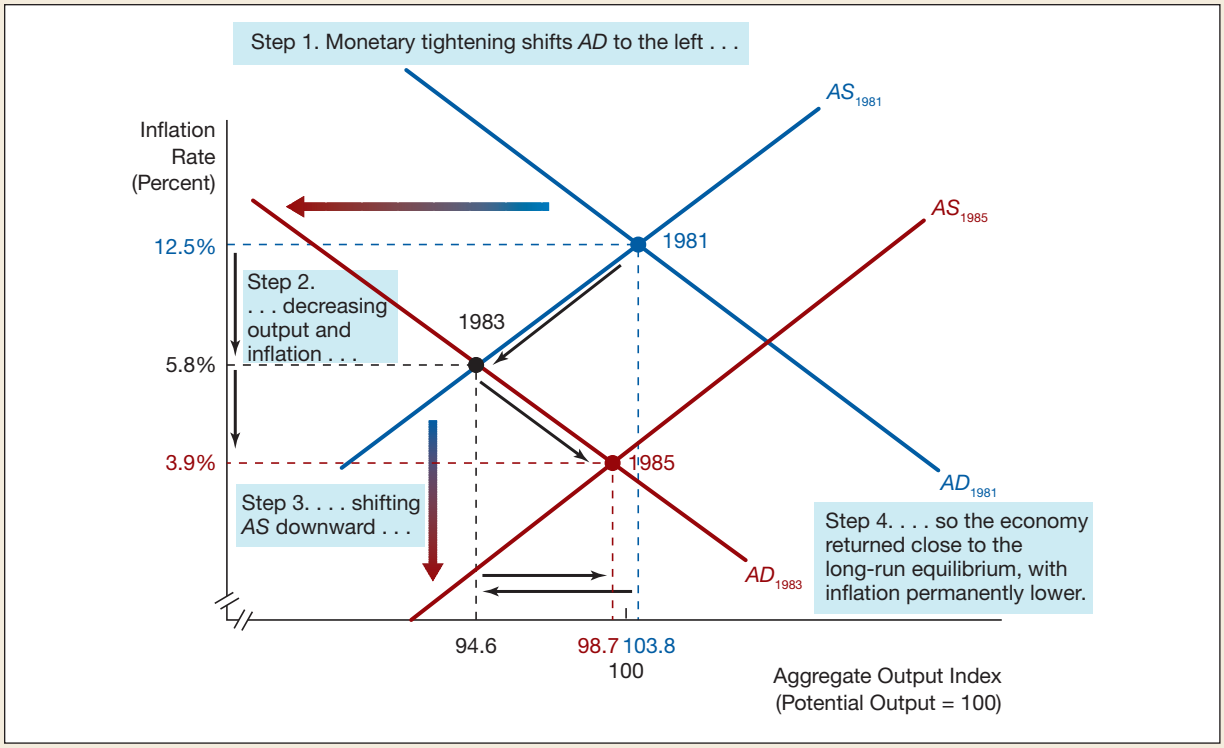


FIGURE 23-14 The Bank of Canada Disinflation

The figure shows that the Bank of Canada actions to decrease inflation were successful but costly: The autonomous monetary policy tightening caused a negative demand shock that shifted the aggregate demand curve to the left from AD_{1981} to AD_{1983} and led to a decline in inflation from 12.5% to 5.8% and a decline in the aggregate output index from 103.8 to 94.6. Then with a further downward shift of the short-run aggregate supply curve to AS_{1985}, the inflation rate fell to 3.9% in 1985, while the aggregate output index recovered to 98.7.

With aggregate output well below its potential output level, the short-run aggregate supply curve shifted downward and to the right to AS_{1985}. By 1985, the economy moved toward the equilibrium at point *1985*. The inflation rate fell further to 3.9% and output rose back near potential output, with the aggregate output index rising to 98.7. ◆

23.7 CHANGES IN EQUILIBRIUM: AGGREGATE SUPPLY (INFLATION) SHOCKS

LO 23.7 Illustrate and interpret the short-run and long-run effects of supply shocks.

In our discussion earlier in the chapter of the short-run aggregate supply curve, we showed that inflation will rise independently of tightness in the labour markets or increases in expected inflation if there are supply shocks, such as a decrease in the supply of oil that causes prices to rise. When the shock involves a restriction in supply,

we refer to this type of supply shock as a *negative* (or *unfavourable*) *supply shock*, and it results in a rise in commodity prices. Examples of negative supply shocks include a disruption in oil supplies, a rise in import prices when a currency declines in value, or a cost-push shock from workers pushing for higher wages that outpace productivity growth, driving up costs and inflation. When the supply shock involves an increase in supply, it is called a *positive* (or *favourable*) *supply shock*. Positive supply shocks include a particularly good harvest or a fall in import prices.

To see how a supply shock affects the economy, we can use our aggregate supply and demand analysis. We start by assuming that output is at its potential level, with the aggregate output index at 100, and inflation is at 2%, at point 1 in Figure 23-15. Suppose that a war in the Middle East causes a negative supply shock. When the negative supply shock hits the economy and oil prices rise, the change in the price shock term ρ indicates that inflation will rise to above 2% and the short-run aggregate supply curve will shift up and to the left, from AS_1 to AS_2.

The economy then will move up the aggregate demand curve, from point 1 to point 2, where inflation rises to 3.5% but the aggregate output index *falls* below 100. We use the term **stagflation** (a combination of the words *stagnation* and *inflation*) to describe

Mini-lecture

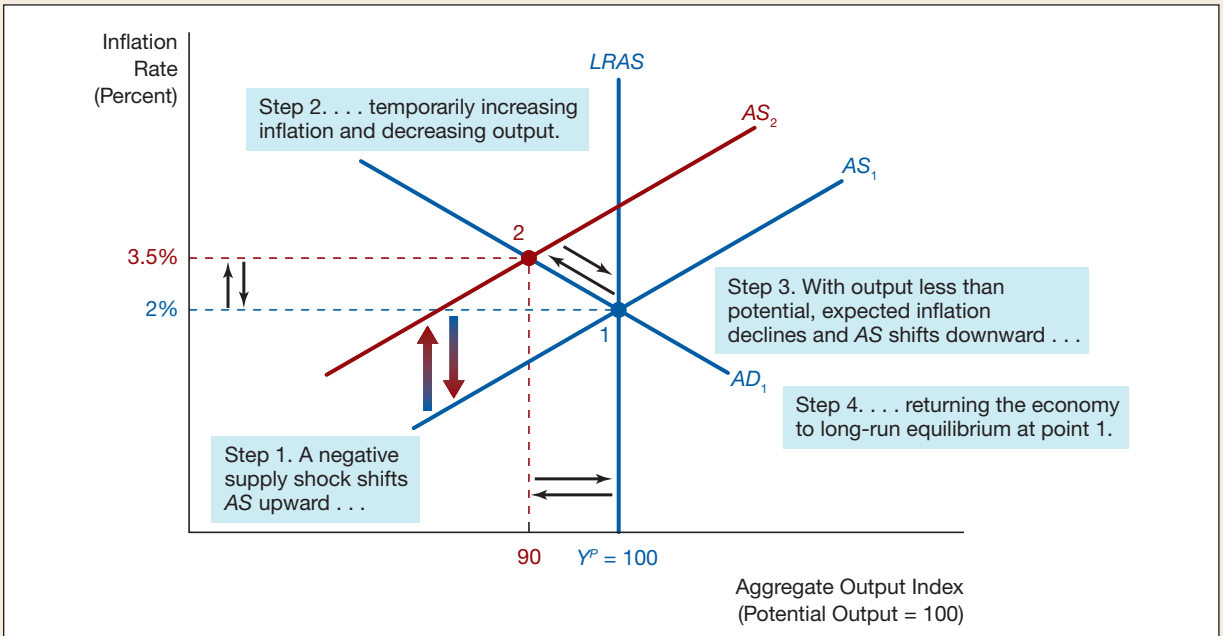

Inflation Rate (Percent)
LRAS
Step 2. . . . temporarily increasing inflation and decreasing output.
AS2
AS1
2
3.5%
2%
1
Step 3. With output less than potential, expected inflation declines and AS shifts downward . . .
AD1
Step 4. . . . returning the economy to long-run equilibrium at point 1.
Step 1. A negative supply shock shifts AS upward . . .
90
YP = 100
Aggregate Output Index (Potential Output = 100)

FIGURE 23-15 Negative Supply Shock

A negative supply shock shifts the short-run aggregate supply curve from AS_1 to AS_2, and the economy moves from point 1 to point 2, where inflation increases to 3.5% and the aggregate output index declines to 90. Because output is less than potential, the negative output gap lowers inflation and hence expected inflation. The decline in expected inflation shifts the short-run aggregate supply curve back down, and eventually it returns to AS_1, where the economy is again at the initial long-run equilibrium at point 1.

a situation defined by rising inflation but a falling level of aggregate output, as pictured in Figure 23-15. At point 2, output is therefore below its potential level (the aggregate output index is, say, at 90), and so inflation falls. The decline in inflation causes expected inflation to fall and shifts the short-run aggregate supply curve back down to where it was initially, at AS_1. The economy (equilibrium) slides down the aggregate demand curve AD_1 (assuming the aggregate demand curve remains in the same position) and returns to the long-run equilibrium at point 1, where the aggregate output index is again at 100 and inflation is again at 2%.

Although a negative supply shock leads to an upward and leftward shift of the short-run aggregate supply curve, which initially raises inflation and lowers output, the ultimate long-run effect is that output and inflation are unchanged.

A favourable (positive) supply shock—say, an excellent harvest of wheat in the Prairie provinces—moves all the curves in Figure 23-15 in the opposite direction and so has the opposite effects. ***A positive supply shock shifts the short-run aggregate supply curve downward and to the right, initially leading to a fall in inflation and a rise in output. In the long run, however, output and inflation are unchanged (holding the aggregate demand curve constant).***

We now will apply the aggregate demand and supply model to studying the negative supply shocks of 1973–1975 and 1978–1980. (Recall we assume that aggregate output is initially at its potential level, with the aggregate output index at 100.)

APPLICATION Negative Supply Shocks in the United States, 1973–1975 and 1978–1980

In 1973, the U.S. economy was hit by a series of negative supply shocks:

1. As a result of the oil embargo stemming from the Arab–Israeli war of 1973, the Organization of Petroleum Exporting Countries engineered a quadrupling of oil prices by restricting oil production.
2. A series of crop failures throughout the world led to a sharp increase in food prices.
3. The termination of U.S. wage and price controls in 1973 and 1974 led to a push by workers to obtain wage increases that had been prevented by the controls.

In 1973, the U.S. inflation rate was 6.3%, while the aggregate output index was at 103.4, as is indicated by the equilibrium point, *1973,* in Figure 23-16. The triple thrust of these negative supply shocks shifted the short-run aggregate supply curve sharply upward and to the left, from AS_{1973} to AS_{1975}, and the economy moved to point *1975.* As shown in Figure 23-16, inflation rose to 9.1% in 1975 and the aggregate output index fell to 95.9. The decline in aggregate output led to a steep rise in the unemployment rate from 4.8% to 8.3%.

The 1978–1980 period was almost an exact replay of the 1973–1975 period. By 1978, the economy had just about fully recovered from the 1973–1975 supply shocks when poor harvests and a doubling of oil prices (resulting from the overthrow of the Shah of Iran) led to another sharp upward and leftward shift of the short-run aggregate supply curve in 1979. The pattern predicted by Figure 23-16 played itself out again—inflation and unemployment both shot upward. ◆

Mini-lecture

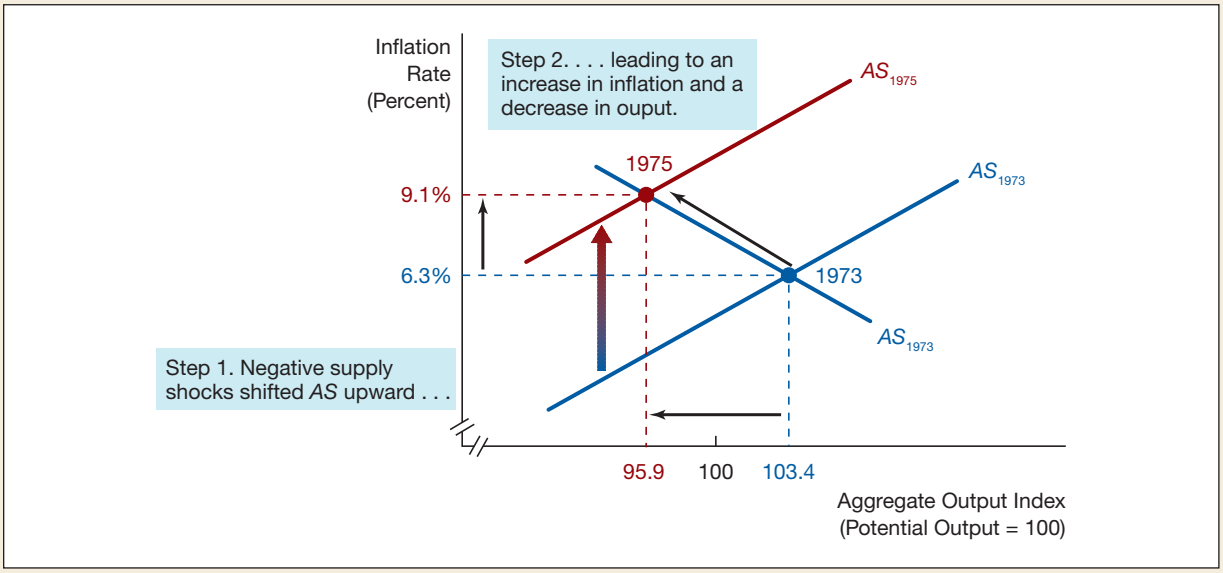


FIGURE 23-16 Negative Supply Shocks in the United States, 1973–1975

The negative supply shocks in 1973 led to an upward shift in the short-run aggregate supply curve, from AS_{1973} to AS_{1975}. The U.S. economy moved from point *1973* to point *1975*, where the aggregate output index fell from 103.4 to 95.9, while the inflation rate rose from 6.3% to 9.1%.

23.8 CONCLUSIONS FROM AGGREGATE DEMAND AND SUPPLY ANALYSIS

LO 23.8 Summarize the conclusions from aggregate demand and supply analysis.

Aggregate demand and supply analysis leads us to the following conclusions.[6]

1. The economy has a self-correcting mechanism that returns it to potential output and the natural rate of unemployment over time.
2. A shift in the aggregate demand curve—caused by changes in autonomous monetary policy (changes in the real interest rate at any given inflation rate), government purchases, taxes, autonomous net exports, autonomous consumption expenditure, autonomous investment, or financial frictions—affects the output gap only in the short run and not in the long run. Furthermore, the initial change in inflation is lower than the long-run change in inflation when the short-run aggregate supply curve has fully adjusted.
3. A supply shock affects the output gap and inflation only in the short run and has no effect in the long run (holding the aggregate demand curve constant).

[6]Aggregate demand and supply analysis also can be used to understand the effects of macroeconomic shocks on asset prices. This analysis can be found in the second online appendix to this chapter.

Often, aggregate demand and aggregate supply shocks occur simultaneously, affecting both the short-run aggregate supply curve and the aggregate demand curve. This is what happened both in the Great Recession of 2007–2009 and in the Covid-19 recession that began in March 2020, as is illustrated in the following applications.

APPLICATION *AD/AS* Analysis of the Great Recession of 2007–2009 in the United States

We break the *AD/AS* analysis of the Great Recession in the United States into two phases. The first phase occurs prior to the bankruptcy of Lehman Brothers, a global financial services firm, in fall 2008, the largest bankruptcy in U.S. history. (We discussed the Lehman Brothers bankruptcy in Chapter 11.)

The second phase occurs from fall 2008 until the end of the recession in 2009. With the bankruptcy of Lehman Brothers in fall 2008, major lenders such as banks stopped lending money, which caused a collapse in investment spending. At the same time, the stock market crashed and consumer optimism plummeted. The result was a sharp cutback in consumer spending.

First Phase. In 2006, the inflation rate in the United States was 2.4% and aggregate output was near potential output, with the aggregate output index at 100.4. This equilibrium at point *2006* is shown in the *AD/AS* diagram in Figure 23-17, which illustrates what happened in the first phase of the Great Recession.

Mini-lecture

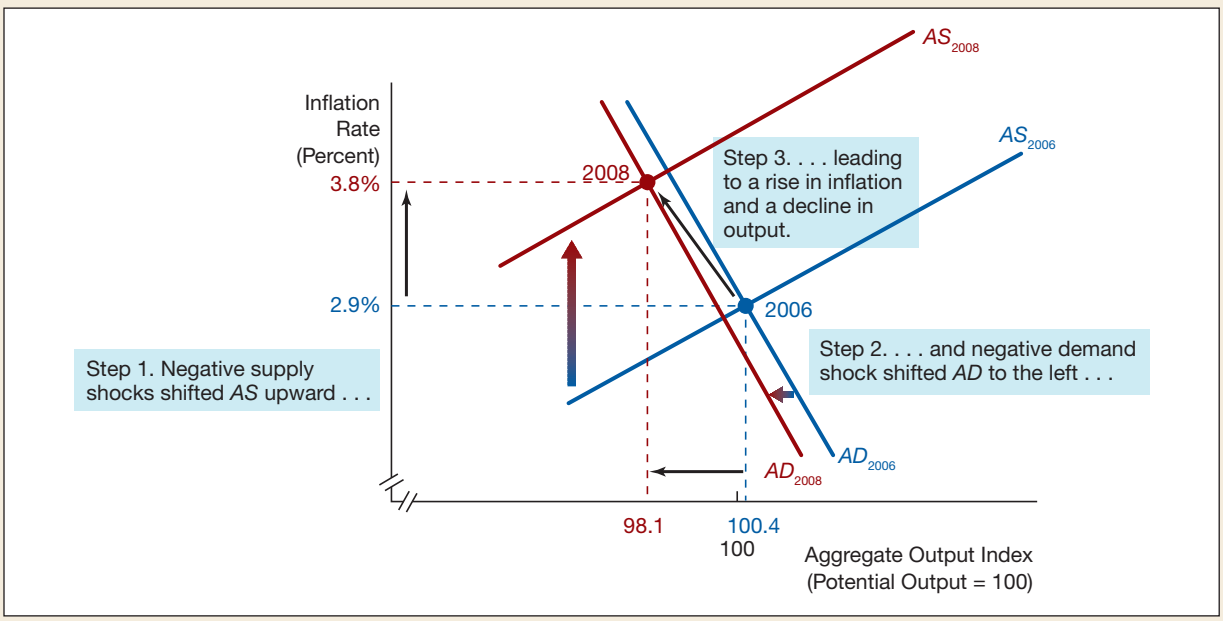

AS_{2008}
AS_{2006}
Inflation Rate (Percent)
3.8%
2.9%
2008
2006
Step 3. . . . leading to a rise in inflation and a decline in output.
Step 1. Negative supply shocks shifted *AS* upward . . .
Step 2. . . . and negative demand shock shifted *AD* to the left . . .
AD_{2006}
AD_{2008}
98.1
100.4
100
Aggregate Output Index (Potential Output = 100)

FIGURE 23-17 First Phase of the Great Recession in the United States

The negative supply shock caused by a rise in the price of oil shifted the short-run aggregate supply curve up from AS_{2006} to AS_{2008}, while the negative demand shock caused by the financial crisis led the aggregate demand curve to shift leftward from AD_{2006} to AD_{2008}. The economy moved to point *2008*, with the aggregate output index falling from 100.4 to 98.1 and the inflation rate rising from 2.9% to 3.8%.

At the beginning of 2007, high demand for oil in rapidly growing developing countries like China and India, and the slowing of oil production in places like Mexico, Russia, and Nigeria, drove oil prices up sharply. Oil prices rose from around US$60 per barrel at the beginning of 2007 to US$100 per barrel at the end of 2007, and they reached a peak of over US$140 per barrel in July 2008. The rise in oil prices, along with increases in other commodity prices, led to a negative supply shock that shifted the short-run aggregate supply curve sharply upward, from AS_{2006} to AS_{2008} in Figure 23-17. To make matters worse, a financial crisis hit the U.S. economy starting in August 2007, causing a sharp increase in financial frictions, which led to contraction in both household and business spending (discussed in Chapter 11). This negative demand shock shifted the aggregate demand curve to the left, from AD_{2006} to AD_{2008}, and moved the economy to point *2008*, with inflation rising to 3.8%, and the aggregate output index falling to 98.1. As predicted by our aggregate demand and supply analysis, this perfect storm of negative shocks led to a recession starting in December 2007, with the unemployment rate rising from 4.6% in 2006 to 5.8% in 2008.

Second Phase. Figure 23-18 provides an *AD/AS* diagram of the second phase of the Great Recession in the United States. The initial equilibrium is at point *2008*, the same

Mini-lecture

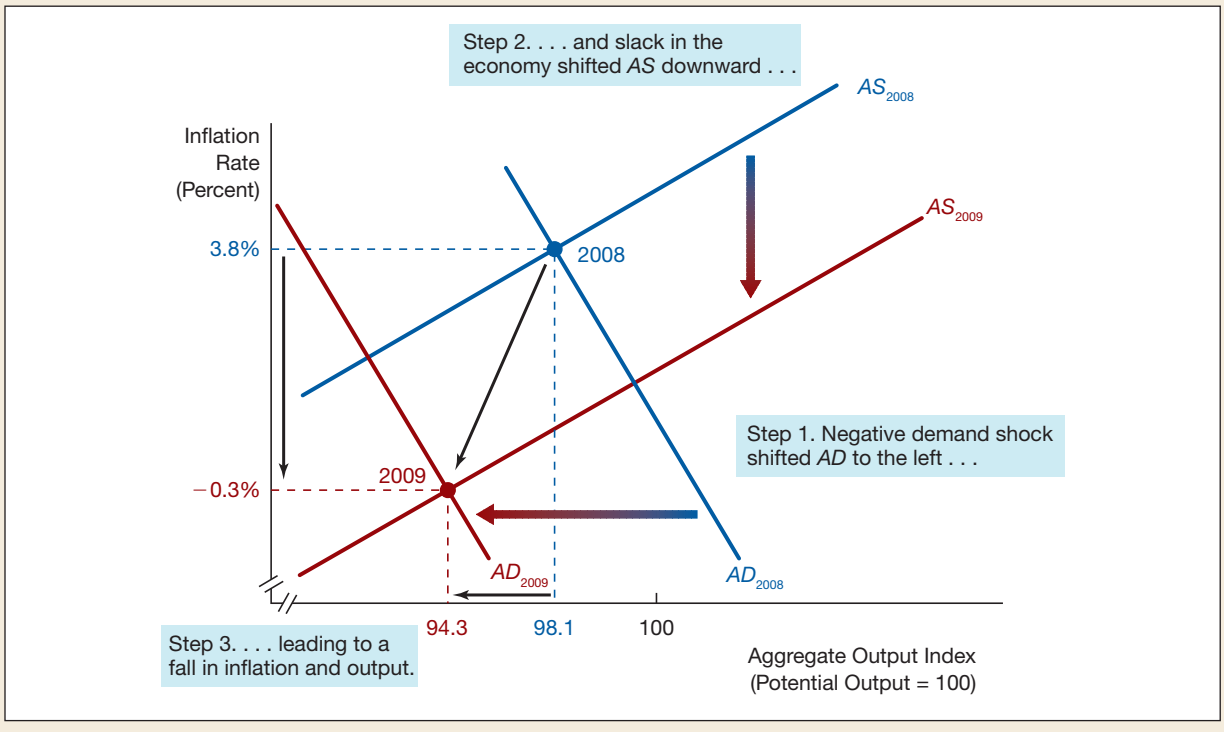


FIGURE 23-18 Second Phase of the Great Recession in the United States

After the Lehman bankruptcy, the worsening financial crisis led to a large leftward shift in the aggregate demand curve from AD_{2008} to AD_{2009}. With commodity prices falling worldwide and aggregate output below potential output, the short-run aggregate supply curve shifted down from AD_{2008} to AS_{2009}. The economy moved to point *2009*, where the inflation rate fell to −0.3% and the aggregate output index fell to 94.3.

as in the previous figure, where the aggregate output index was at 98.1 and the inflation rate was 3.8%. The second phase of the Great Recession began in fall 2008, with the bankruptcy of Lehman Brothers. The shockwave that hit the financial system led to a massive contraction of lending, which caused a collapse in investment spending. At the same time, the stock market crashed, with stock prices falling by more than 50%. Consumer optimism plummeted. The result was a sharp cutback in consumer spending, which shifted the aggregate demand curve sharply to the left from AD_{2008} to AD_{2009}.

As the financial crisis spread to the rest of the world, commodity prices plummeted. In particular, the price of oil fell to $40 per barrel by early 2009, well below the $60 price at the beginning of 2007. This decline in oil prices, and the fact that aggregate output was below its potential level, led to a sharp decline in the short-run aggregate supply curve to AS_{2009}. As shown in Figure 23-18, the equilibrium moved to point *2009*, where the aggregate output index fell to 94.3 and inflation fell to −0.3%. The large decline in aggregate output led to a rise in the unemployment rate to 9.3% in 2009, but in the aftermath of the financial crisis, the economic recovery was very slow, with aggregate output not returning to its potential level until 10 years after the financial crisis started (see Figure 23-3). This is why the recession of 2007–2009 has been given the name the Great Recession. ◆

APPLICATION An *AD/AS* Analysis of the Covid-19 Recession in Canada

When the coronavirus began to spread exponentially in March 2020 and lockdowns were implemented throughout Canada, the economy went into a tailspin. The economy declined at the steepest rate in Canadian history, with the aggregate output index falling from 99.2 in the fourth quarter of 2019 to 85.4 in the second quarter of 2020. Note that in this application, we use one of the Bank of Canada's estimates of the output gap (based on the integrated framework). With this precipitous drop in output, the unemployment rate rose from 5.7% in the fourth quarter of 2019 to 13% in the second quarter of 2020. At the same time, the inflation rate fell from 2.1% to (near) 0.0%.

What is unusual about the resulting Covid-19 recession is that it was triggered by a massive aggregate supply shock that induced a massive aggregate demand shock. The lockdowns that began in March 2020 led to production shutting down in many parts of the Canadian economy. The result was that the short-run aggregate supply curve shifted up and to the left from $AS_{2019:Q4}$ to $AS_{2020:Q2}$ in Figure 23-19. Because the lockdowns meant that households were unable to venture out to stores, they sharply cut back their spending. In addition, because it was hard to know how long the coronavirus pandemic would last and when the economy might reopen, uncertainty increased greatly, so businesses shelved a lot of their investment plans. Uncertainty and the collapse in stock prices also further reduced the willingness of consumers to spend. The result was a large decline in the quantity of aggregate output demanded at any given inflation rate, so the aggregate demand curve shifted sharply to the left from $AD_{2019:Q4}$ to $AD_{2020:Q2}$.

Mini-lecture

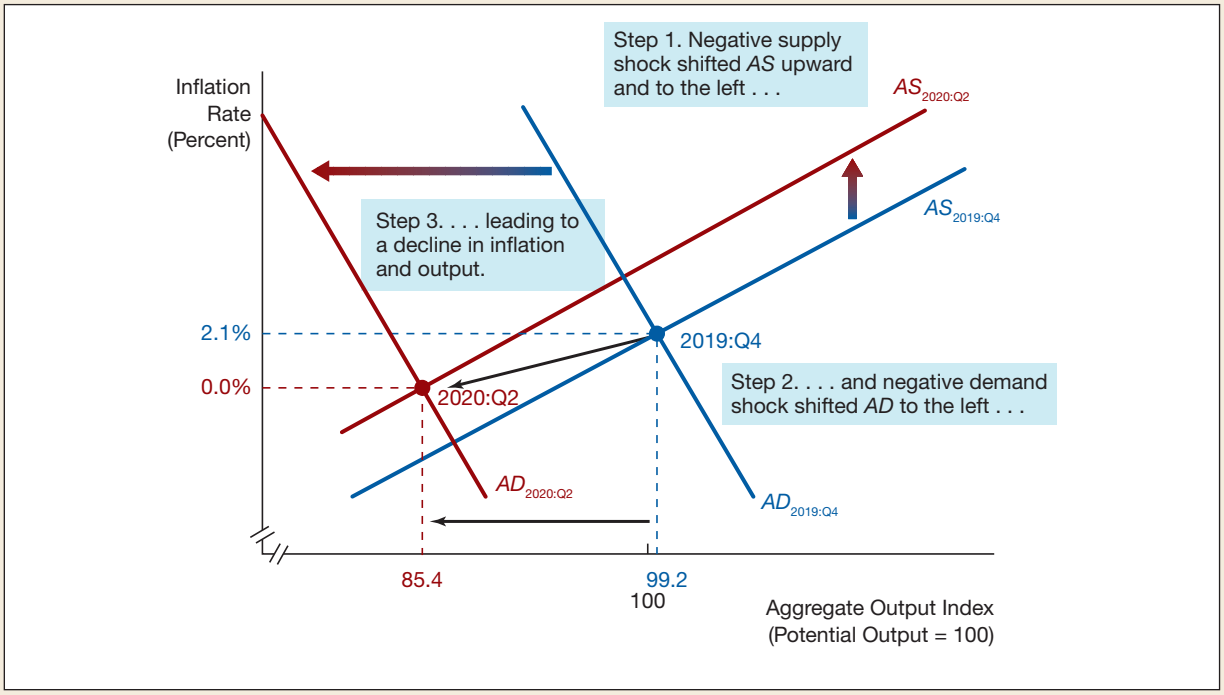

Step 1. Negative supply shock shifted AS upward and to the left . . .
$AS_{2020:Q2}$
$AS_{2019:Q4}$
Inflation Rate (Percent)
Step 3. . . . leading to a decline in inflation and output.
2.1%
0.0%
2019:Q4
2020:Q2
Step 2. . . . and negative demand shock shifted AD to the left . . .
$AD_{2020:Q2}$
$AD_{2019:Q4}$
85.4
99.2
100
Aggregate Output Index (Potential Output = 100)

FIGURE 23-19 The Covid-19 Recession in Canada

The lockdown of the Canadian economy led to the short-run aggregate supply curve shifting from $AS_{2019:Q4}$ to $AS_{2020:Q2}$, which also led to a leftward shift of the aggregate demand curve from $AD_{2019:Q4}$ to $AD_{2020:Q2}$. The macroeconomic equilibrium moved from point *2019:Q4* to point *2020:Q2*, and the aggregate output index fell to 85.4 and inflation fell to 0.0%.

The new equilibrium moved from point *2019:Q4* to point *2020:Q2*. The aggregate output index fell from 99.2 in the 2019:Q4 to 85.4 in 2020:Q2, with the unemployment rate rising to 13%. Although our analysis predicts that aggregate output would necessarily fall because both the aggregate demand and short-run aggregate supply curves shifted to the left, it was not clear whether inflation would rise or fall because the restriction in supply would raise inflation, while the decline in aggregate demand would lower inflation. However, another event that occurred was a price war between Saudi Arabia and Russia, which led to a drop in the price of oil from around US$60 a barrel to less than US$30 per barrel in the second quarter of 2020. This precipitous drop in the price of oil acted as a positive supply shock that offset some of the negative supply shock from the pandemic. As a result, the short-run aggregate supply curve shifted by less than the aggregate demand curve, and the inflation rate fell from 2.1% in the fourth quarter of 2019 to 0.0% in the second quarter of 2020. ◆

SUMMARY

1. Business cycles are the upward and downward movement of aggregate economic activity and are important for aggregate demand and supply analysis to explain because declines in aggregate output are associated with increases in unemployment, which lead to economic hardship. Aggregate demand and supply analysis is also used to explain fluctuations in inflation, which are important because low and stable inflation reduces the need for businesses to engage in costly changes in price for their products and reduces uncertainty in the economy, making it easier for households and businesses to make economic decisions.
2. The aggregate demand curve indicates the quantity of aggregate output demanded at each inflation rate, and it is downward-sloping. The primary sources of shifts in the aggregate demand curve are (1) autonomous monetary policy, (2) government purchases, (3) taxes, (4) autonomous net exports, (5) autonomous consumption expenditure, (6) autonomous investment, and (7) financial frictions.
3. The long-run aggregate supply curve is vertical at potential output. The long-run aggregate supply curve shifts when technology changes, when there are long-run changes to the amount of labour or capital, or when the natural rate of unemployment changes. The short-run aggregate supply curve slopes upward because inflation rises as output rises relative to potential output. The short-run supply curve shifts when there are changes in expected inflation, inflation shocks, or persistent output gaps.
4. Equilibrium in the short run occurs at the point at which the aggregate demand curve intersects the short-run aggregate supply curve. Although this is the level to which the economy heads temporarily, the self-correcting mechanism leads the economy to settle permanently at the long-run equilibrium, the point at which the aggregate demand curve intersects the long-run aggregate supply curve and aggregate output is at its potential. Shifts in either the aggregate demand curve or the short-run or long-run aggregate supply curves can produce changes in aggregate output and inflation.
5. A positive demand shock shifts the aggregate demand curve to the right and initially leads to a rise in both inflation and output. However, in the long run, such a shock leads only to a rise in inflation, because output returns to its potential initial level.
6. A positive supply shock leads to a downward and rightward shift in the short-run aggregate supply curve, which initially lowers inflation and raises output. However, in the long run, the output gap and inflation are unchanged.
7. Three conclusions can be drawn from aggregate demand and supply analysis: (1) The economy has a self-correcting mechanism that returns it to potential output and the natural rate of unemployment over time; (2) a shift in the aggregate demand curve—caused by changes in autonomous monetary policy (changes in the real interest rate at any given inflation rate), government purchases, taxes, autonomous net exports, autonomous consumption expenditure, autonomous investment, or financial frictions—affects the output gap only in the short run and not in the long run; furthermore, the initial change in inflation is lower than the change in the long run; and (3) a supply shock affects the output gap and inflation only in the short run and has no effect in the long run (holding the aggregate demand curve constant).

MyLab Economics Make the grade with MyLab Economics: Select end-of-chapter questions can be found on MyLab Economics. You can practise them as often as you want, and most feature step-by-step guided instructions to help you find the right answer.

QUESTIONS

1. Explain why the aggregate demand curve slopes downward and the short-run aggregate supply curve slopes upward.
2. Identify three factors that can shift the aggregate demand curve to the right and three different factors that can shift the aggregate demand curve to the left.

3. "The appreciation of the U.S. dollar from 2012 to 2018 had a negative effect on aggregate demand in the United States." Is this statement true, false, or uncertain? Explain your answer.
4. In many countries around the world, the population is aging and large segments of the population are retiring or close to retirement. What effect would this have on a country's long-run aggregate supply curve? What will happen to aggregate output as a result?
5. If the labour force becomes more productive over time, how would the long-run aggregate supply curve be affected?
6. Why are central banks so concerned with inflation expectations?
7. "If prices and wages are perfectly flexible, then $\gamma = 0$ and changes in aggregate demand have a smaller effect on output." Is this statement true, false, or uncertain? Explain your answer.
8. What factors shift the short-run aggregate supply curve? Do any of these factors shift the long-run aggregate supply curve? Why?
9. If large budget deficits cause the public to think there will be higher inflation in the future, what is likely to happen to the short-run aggregate supply curve when budget deficits rise?
10. In the aftermath of the financial crisis in the United States, labour mobility has decreased significantly. How, if at all, might this affect the natural rate of unemployment?
11. When aggregate output is below the natural rate of output, what happens to the inflation rate over time if the aggregate demand curve remains unchanged? Why?
12. Suppose the public believes that a newly announced anti-inflation program will work and so lowers its expectations of future inflation. What will happen to aggregate output and the inflation rate in the short run?
13. If the unemployment rate is above the natural rate of unemployment, holding other factors constant, what will happen to inflation and output?
14. What happens to inflation and output in the short run and the long run when taxes decrease?
15. If stagflation is bad (high inflation and high unemployment), does this necessarily mean low inflation and low unemployment is good?
16. Why did the Bank of Canada pursue inherently recessionary policies in the early 1980s?
17. In what ways is the Bank of Canada's disinflation considered a success? In what ways is it considered a failure?

APPLIED PROBLEMS

18. Suppose the government passes legislation that encourages investment in research and the development of new technologies. Assuming this policy leads to a positive productivity change for the Canadian economy, use aggregate demand and supply analysis to predict the effects on inflation and output. Demonstrate these effects on a graph.
19. Proposals have been made that advocate an increase in the national sales tax. Predict the effects of such a tax on the aggregate supply and demand curves, showing the effects on output and inflation. Use a graph of aggregate supply and demand to demonstrate these effects.
20. Suppose the inflation rate remains relatively constant while output decreases and the unemployment rate increases. Using an aggregate demand and supply graph, show how this scenario is possible.
21. Classify each of the following as a supply shock or a demand shock. Use a graph to show the effects on inflation and output in the short run and in the long run.
    a. Financial frictions increase.
    b. Households and firms become more optimistic about the economy.
    c. Favourable weather produces a record crop of wheat and corn in the Prairie provinces.
    d. Auto workers go on strike for four months.
22. During 2017, some Bank of Canada officials discussed the possibility of increasing interest rates as a way of fighting potential increases in expected inflation. If the public came to expect higher inflation rates in the future, what would be the effect on the short-run aggregate supply curve? Use an aggregate demand and supply graph to illustrate your answer.

DATA ANALYSIS PROBLEMS

1. Go to the CANSIM database and find monthly data from January 1992 to December 2020 on the total unemployment rate (series V2062815) and the (all items) consumer price index (series V41690914).
    a. What is the average annual inflation rate?
    b. Is the correlation between the annual inflation rate and the unemployment rate negative or positive?
2. **Real-Time Data Analysis** Go to the St. Louis Federal Reserve FRED database and find U.S. data on real GDP (GDPC1), potential GDP (GDPPOT), and the unemployment rate (UNRATE) from 1960 to the most recent period. For the unemployment rate data, convert the frequency to "Quarterly." Download all of the data into a spreadsheet, and calculate the output gap and aggregate output index for each quarter.
    a. In which quarter is the output gap the most negative? What is the Aggregate Output Index for that quarter?
    b. In which quarter is the unemployment rate highest?
    c. Are these periods of time close to each other? What does this tell you about the relationship between output gaps and the unemployment rate?
3. **Real-Time Data Analysis** Go to the St. Louis Federal Reserve FRED database and find data on real government spending (GCEC1), real GDP (GDPC1), taxes (W006RC1Q027SBEA), and the personal consumption expenditure price index (PCECTPI), a measure of the price level. Download all of the data into a spreadsheet, and convert the tax data series into real taxes. To do this, for each quarter, divide taxes by the price index and then multiply by 100.
    a. Calculate the level change in real GDP over the four most recent quarters of data available and the four quarters prior to that.
    b. Calculate the level change in real government spending and real taxes over the four most recent quarters of data available and the four quarters prior to that.
    c. Are your results consistent with what you would expect? How do your answers to part (b) help explain, if at all, your answer to part (a)? Explain using the *IS* and *AD* curves.

24 Monetary Policy Theory

Learning Objectives

24.1 Illustrate and explain the policy choices that monetary policymakers face under the conditions of aggregate demand and aggregate supply shocks.

24.2 Identify the lags in the policy process, and summarize why they weaken the case for an activist policy approach.

24.3 Explain why monetary policymakers can target any inflation rate in the long run but cannot target a level of aggregate output in the long run.

24.4 Identify the sources of inflation and the role of monetary policy in propagating inflation.

24.5 Explain the unique challenges that monetary policymakers face at the effective lower bound, and illustrate how nonconventional monetary policy can be effective under such conditions.

Preview

Between November 2007 and April 2009, the Bank of Canada lowered the target for its policy interest rate, the overnight interest rate, from 4.5% all the way down to 0.25% and continued to keep it around 1% for more than five years afterward. Why did the Bank of Canada lower interest rates this aggressively and continue to keep them so low?

Could this monetary policy easing have sparked undesirable inflation? Many commentators in the media thought so. Starting in the early 1960s, when the inflation rate hovered between 1% and 2%, the economy began to suffer from higher and more variable rates of inflation. By 1974, it reached the double-digit level. After moderating somewhat during the 1976–1978 period, it shot close to 13% in 1981, slowed to around 4% from 1984 to 1988, declined further to less than 2% in the early 1990s, and then climbed to above 4% in 2003. Subsequently, the inflation rate fell to around the 2% level. Inflation has become a major concern of politicians and the public, and how best to control it is an issue that frequently dominates the discussion of economic policy.

In this chapter, we will use the aggregate demand/aggregate supply (*AD/AS*) framework developed in Chapter 23 to develop a theory of monetary policy. Specifically, we will examine the role of monetary policy in stabilizing the economy and creating inflation. We apply the theory to four big questions: Does stabilizing inflation stabilize output? Should policy be *activist*, responding aggressively to fluctuations in economic activity, or passive and *nonactivist*? What are the roots of inflation? How can monetary policy work when interest rates cannot go much below zero?

24.1 RESPONSE OF MONETARY POLICY TO SHOCKS

LO 24.1 Illustrate and explain the policy choices that monetary policymakers face under the conditions of aggregate demand and aggregate supply shocks.

As we saw in Chapter 17, the primary goal of central banks is price stability: That is, they try to maintain inflation, π, close to a target level (π^T), referred to as an **inflation target**, that is slightly above zero. Most central banks set π^T between 1% and 3%. In other words, central banks pursue price stability through monetary policy that aims to minimize the difference between inflation and the inflation target ($\pi - \pi^T$), a difference economists refer to as the **inflation gap**.

In Chapter 17, we learned that central banks also care about stabilizing economic activity. Because economic activity can be sustained only at potential output, we can describe this objective of monetary policy by saying that monetary policymakers want aggregate output to be close to its potential level, Y^P. Central banks want to minimize the difference between aggregate output and potential output ($Y - Y^P$), in other words, the *output gap*. In our analysis of aggregate demand and supply in Chapter 23, we examined two categories of economic shocks—demand shocks and supply shocks—and the effects of each on inflation and output. In this section, we describe a central bank's policy responses, given its objectives, to each of these types of shocks. In the case of demand shocks, policymakers can simultaneously pursue price stability and stability of economic activity. In the aftermath of a supply shock, however, policymakers can achieve either price stability or economic activity stability but not both. This trade-off poses a thorny dilemma for central banks with dual mandates.

Response to an Aggregate Demand Shock

Because we are discussing stabilizing economic activity, we need to focus on output gaps rather than the level of aggregate output. Thus, in this and the next chapter, we will conduct all *AD/AS* analysis in terms of the aggregate output index, in which potential output has a value of 100, rather than in terms of aggregate output. In all the following *AD/AS* diagrams, the label on the horizontal axis is now always going to be the aggregate output index, and Y represents the aggregate output index rather than aggregate output.

We begin by considering the effects of an aggregate demand shock, illustrated in Figure 24-1. For example, the coronavirus pandemic began, in March 2020, to cause both consumer and business spending to fall. In the figure, the economy is initially at point 1, where output is at its potential level so the aggregate output index is at $Y^P = 100$, and inflation is at π^T. The negative demand shock decreases aggregate demand, shifting AD_1 to the left to AD_2. Policymakers can respond to this shock in two possible ways.

No Policy Response If the central bank does not respond by changing the autonomous component of monetary policy, the aggregate demand curve remains at AD_2, and so the economy moves to the intersection of AS_1 and AD_2. Here, the aggregate output index falls to Y_2, so aggregate output is below potential output, and inflation falls to π_2, below the inflation target of π^T. With inflation falling and output below potential, expected inflation will decline and the short-run aggregate supply curve will shift down and to the right until it reaches AS_3, and the economy will move to point 3. Output will be back to its potential level at an aggregate output index of $Y^P = 100$ and inflation will fall to a lower level of π_3. At first glance, this outcome looks favourable—inflation is lower and output is back at its potential. But aggregate output will remain below potential for some time, and if inflation was initially at its target level, the fall in inflation is undesirable for the reasons outlined in Chapter 17.

Policy Stabilizes Economic Activity and Inflation in the Short Run Policymakers can eliminate both the output gap and the inflation gap in the short run by pursuing policies aimed at increasing aggregate demand to its initial level and returning the economy to its preshock state. The central bank does this by autonomously easing monetary policy by cutting the real interest rate at any given inflation rate. This action stimulates investment spending and increases the quantity of aggregate output demanded at any given inflation rate, thereby shifting the *AD* curve to the right. As a result, the aggregate demand curve shifts from AD_2 back to AD_1 in Figure 24-2,

Mini-lecture

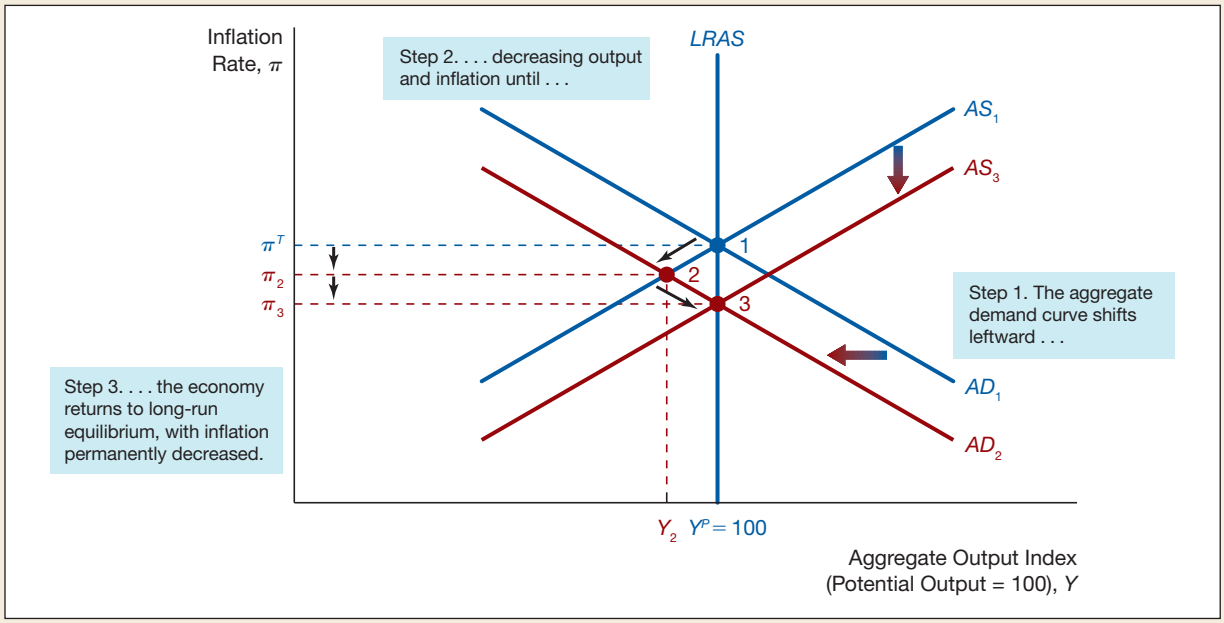


FIGURE 24-1 Aggregate Demand Shock: No Policy Response

An aggregate demand shock shifts the aggregate demand curve leftward from AD_1 to AD_2 and moves the economy from point 1 to point 2, where the aggregate output index falls to Y_2 and inflation falls to π_2. With output below potential, the short-run aggregate supply curve shifts down to AS_3 and the economy moves to point 3, where output is back at $Y^P = 100$ but inflation has fallen to π_3.

and the economy returns to point 1. (The Federal Reserve in the United States took exactly these steps by lowering the federal funds rate from 5% to ¼% to zero over 15 months starting in September 2007.)

Our analysis of this monetary policy response shows that ***in the case of aggregate demand shocks, there is no trade-off between the pursuits of price stability and economic activity stability.*** A monetary policy response that focuses on stabilizing inflation is exactly the monetary policy response that will stabilize economic activity. No conflict exists between the dual objectives of stabilizing inflation and stabilizing economic activity, a result that Olivier Blanchard (the chief economist at the International Monetary Fund at the time, formerly a professor at MIT) referred to as the **divine coincidence**.

Response to a Supply Shock

When a supply shock occurs—for example, when the price of oil surges because of political unrest in the Middle East or because of an act of God such as a devastating drought in the Prairies—the divine coincidence does not hold. Policymakers face a short-run trade-off between stabilizing inflation and stabilizing economic activity. To illustrate, we start with the economy at point 1 in Figure 24-3, where aggregate output is at the natural rate, so the aggregate output index is at $Y^P = 100$ and inflation is at π^T.

Mini-lecture

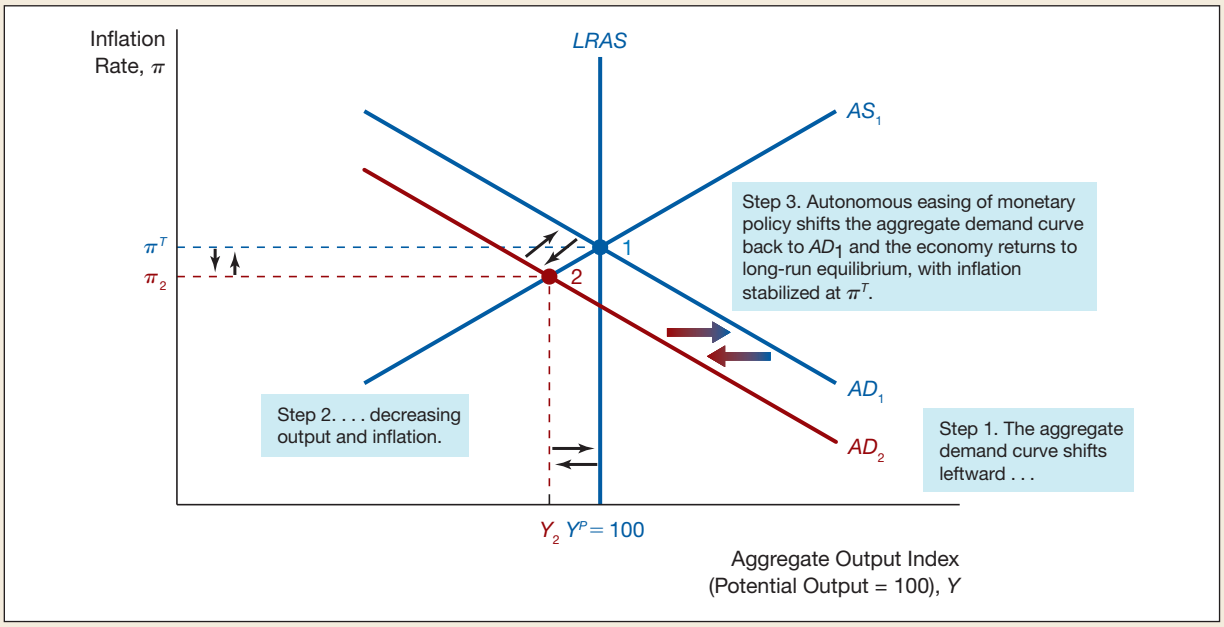

Inflation Rate, π
LRAS
AS1
Step 3. Autonomous easing of monetary policy shifts the aggregate demand curve back to AD1 and the economy returns to long-run equilibrium, with inflation stabilized at π^T.
π^T
π2
1
2
AD1
AD2
Step 2. . . . decreasing output and inflation.
Step 1. The aggregate demand curve shifts leftward . . .
Y2 Y^P = 100
Aggregate Output Index (Potential Output = 100), Y

FIGURE 24-2 Aggregate Demand Shock: Policy Stabilizes Output and Inflation in the Short Run

An aggregate demand shock shifts the aggregate demand curve leftward from AD_1 to AD_2 and moves the economy from point 1 to point 2, where the aggregate output index falls to Y_2 and inflation falls to π_2. An autonomous easing of monetary policy lowers the real interest rate at any given inflation rate and shifts the *AD* curve back to AD_1. Aggregate output returns to potential at point 1, and inflation returns to its target level.

The negative supply shock—from, say, a rise in the price of oil—shifts the short-run aggregate supply curve up and to the left, from AS_1 to AS_2.The economy moves to point 2, with inflation rising to π_2 and the aggregate output index falling to Y_2. Policymakers can respond to the supply shock in three possible ways.

No Policy Response One potential policy choice is to refrain from making an autonomous change in monetary policy so that the aggregate demand curve does not shift. Since aggregate output is less than potential output, eventually the short-run aggregate supply curve will shift back down to the right, returning to AS_1. The economy will return to point 1 in Figure 24-3, and both the output and inflation gaps will close as output and inflation return to their initial levels of $Y^P = 100$ and π^T, respectively. Both inflation and economic activity will stabilize over time. While we wait for the long run, however, the economy will undergo a painful period of reduced output and higher inflation rates. To avoid this period of reduced output and high inflation, monetary policymakers might decide to try to stabilize economic activity or inflation in the short run.

Policy Stabilizes Inflation in the Short Run A second policy choice for monetary authorities is to keep inflation at the target level of π^T in the short run by autonomously tightening monetary policy and raising the real interest rate at any given inflation rate. Doing so will cause investment spending and aggregate demand to fall at

Mini-lecture

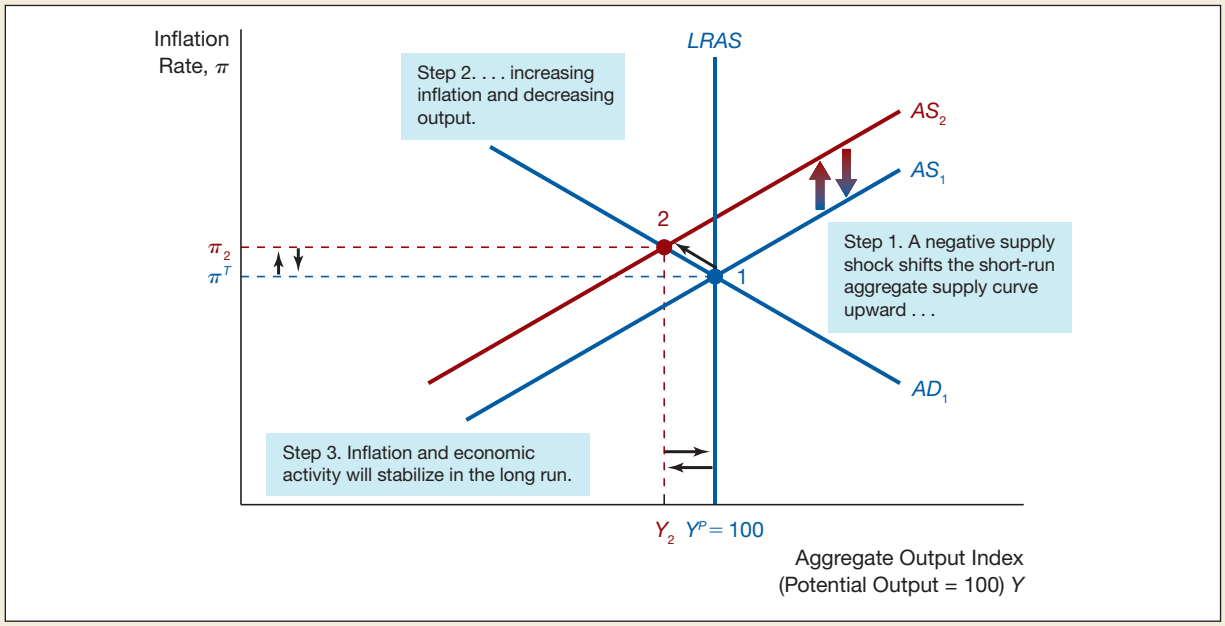

Inflation Rate, π
LRAS
Step 2. . . . increasing inflation and decreasing output.
AS2
AS1
2
π2
πT
1
Step 1. A negative supply shock shifts the short-run aggregate supply curve upward . . .
AD1
Step 3. Inflation and economic activity will stabilize in the long run.
Y2 YP = 100
Aggregate Output Index (Potential Output = 100) Y

FIGURE 24-3 Response to a Supply Shock: No Policy Response

A negative supply shock shifts the short-run aggregate supply curve upward from AS_1 to AS_2, moving the economy to point 2, with inflation rising to π_2 and the aggregate output index falling to Y_2. If the autonomous monetary policy remains unchanged, the short-run aggregate supply curve will shift back down and to the right in the long run, eventually returning to AS_1, and the economy will move back to point 1.

each inflation rate, shifting the aggregate demand curve to the left to AD_3 in Figure 24-4. The economy now will move to point 3, where the aggregate demand curve AD_3 intersects the short-run aggregate supply curve AS_2 at an inflation rate of π^T. Because output is below potential at point 3, the short-run aggregate supply curve will shift back down to AS_1. To keep the inflation rate at π^T, the monetary authorities will need to move the short-run aggregate demand curve back to AD_1 by reversing the autonomous tightening, and eventually the economy will return to point 1.

As Figure 24-4 illustrates, stabilizing inflation reduces the aggregate output index to Y_3 in the short run, and only over time will output return to potential output at an aggregate output index of $Y^P = 100$. ***Stabilizing inflation in response to a supply shock leads to a larger deviation of aggregate output from potential, so this action does not stabilize economic activity.***

Policy Stabilizes Economic Activity in the Short Run A third policy choice is for monetary policymakers to stabilize economic activity rather than inflation in the short run by increasing aggregate demand. According to Figure 24-5, this policy decision will shift the aggregate demand curve to the right to AD_3, where it intersects the short-run aggregate supply curve AS_2 and the long-run aggregate supply curve *LRAS* at point 3. To achieve this result, policymakers will have to autonomously ease monetary policy by lowering the real interest rate at any given inflation rate. At point 3, the output gap

Mini-lecture

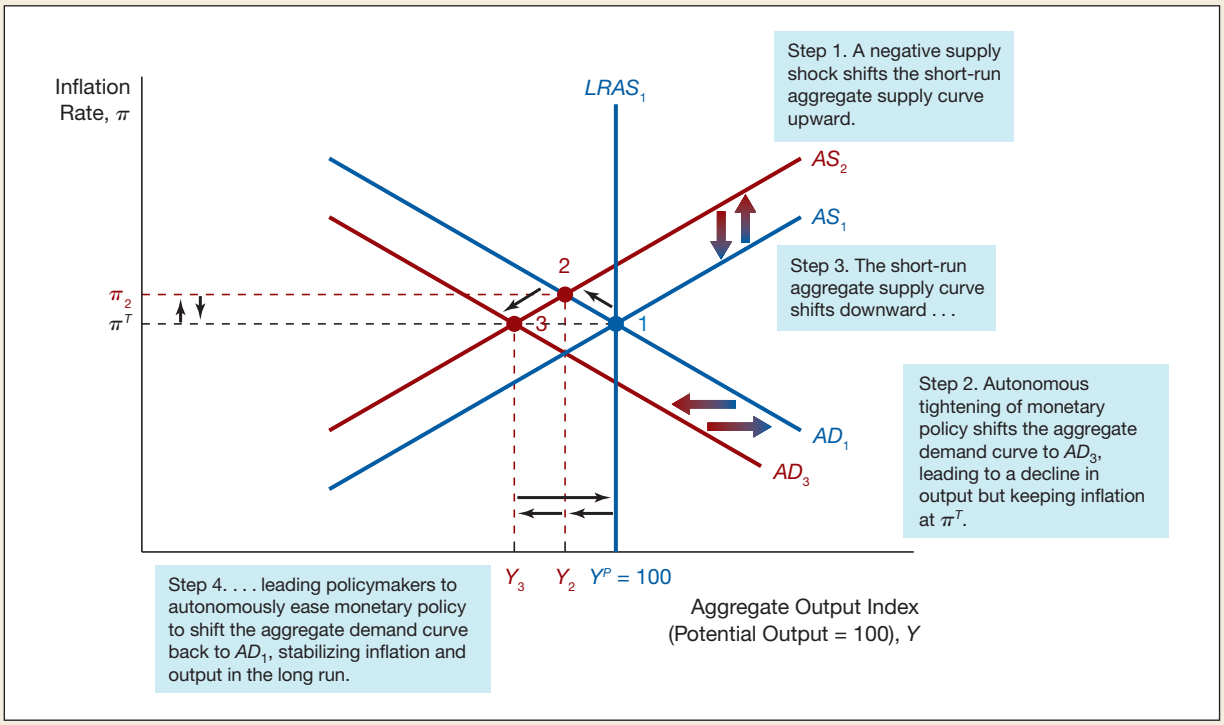

Inflation Rate, π
LRAS1
AS2
AS1
AD1
AD3
π2
πT
1
2
3
Y3
Y2
YP = 100
Aggregate Output Index (Potential Output = 100), Y
Step 1. A negative supply shock shifts the short-run aggregate supply curve upward.
Step 2. Autonomous tightening of monetary policy shifts the aggregate demand curve to AD3, leading to a decline in output but keeping inflation at πT.
Step 3. The short-run aggregate supply curve shifts downward . . .
Step 4. . . . leading policymakers to autonomously ease monetary policy to shift the aggregate demand curve back to AD1, stabilizing inflation and output in the long run.

FIGURE 24-4 Response to a Supply Shock: Short-Run Inflation Stabilization

A negative supply shock shifts the short-run aggregate supply curve from AS_1 to AS_2, moving the economy to point 2, with inflation rising to π_2 and the aggregate output index falling to Y_2. Autonomous tightening of monetary policy shifts the aggregate demand curve to the left to AD_3, and the economy moves to point 3, where inflation is at π^T. With output below potential at point 3, the short-run aggregate supply curve shifts back to AS_1. To keep the inflation rate at π^T, the autonomous tightening of monetary policy is reversed, shifting the aggregate demand curve back to AD_1 and the economy back to point 1.

returns to zero, so monetary policy has stabilized economic activity. However, inflation has risen to π_3, which is greater than π^T, and so inflation has *not* been stabilized. ***Stabilizing economic activity in response to a supply shock results in a larger deviation of inflation from the inflation target rather than a stabilization of inflation.***

The Bottom Line: The Relationship Between Stabilizing Inflation and Stabilizing Economic Activity

We can draw the following conclusions from our analysis:

1. ***If most shocks to the economy are aggregate demand shocks, then policy that stabilizes inflation will also stabilize economic activity, even in the short run.***
2. ***If supply shocks are the most common type of shock, then a central bank must choose between the two stabilization objectives in the short run.***

Mini-lecture

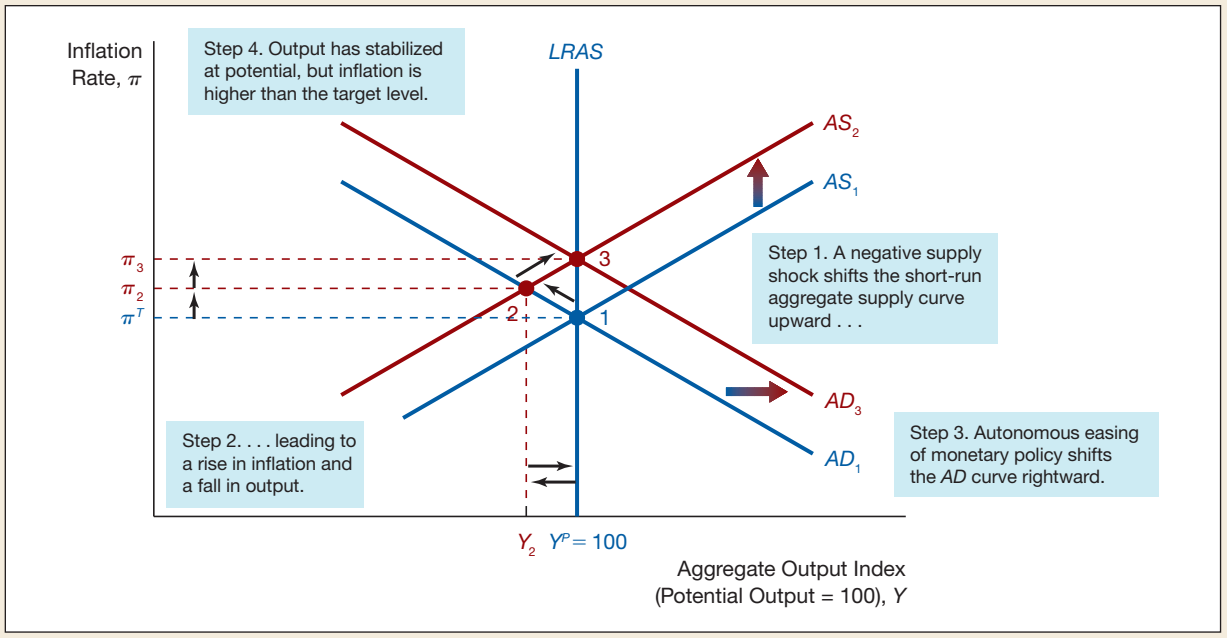

Inflation Rate, π
Step 4. Output has stabilized at potential, but inflation is higher than the target level.
LRAS
AS2
AS1
Step 1. A negative supply shock shifts the short-run aggregate supply curve upward . . .
π3
π2
πT
3
2
1
AD3
AD1
Step 2. . . . leading to a rise in inflation and a fall in output.
Step 3. Autonomous easing of monetary policy shifts the AD curve rightward.
Y2
YP = 100
Aggregate Output Index (Potential Output = 100), Y

FIGURE 24-5 Response to a Supply Shock: Short-Run Output Stabilization

A negative supply shock shifts the short-run aggregate supply curve from AS_1 to AS_2, moving the economy to point 2, with inflation rising to π_2 and the aggregate output index falling to Y_2. To stabilize output, autonomous monetary policy easing shifts the aggregate demand curve rightward to AD_3. At point 3, the monetary policy action has stabilized economic activity, but inflation at π_3 is greater than the target level π^T.

24.2 HOW ACTIVELY SHOULD POLICYMAKERS TRY TO STABILIZE ECONOMIC ACTIVITY?

LO 24.2 Identify the lags in the policy process, and summarize why they weaken the case for an activist policy approach.

All economists have similar policy goals (to promote high employment and price stability), yet they often disagree on the best approach for achieving these goals. Suppose policymakers confront an economy that has high unemployment resulting from a negative demand or supply shock that has reduced aggregate output. **Nonactivists** believe that wages and prices are very flexible and that the self-correcting mechanism works very rapidly. They argue that the short-run aggregate supply curve will shift downward quickly, returning the economy to full employment rapidly. They thus believe that government action to eliminate unemployment is unnecessary. **Activists**, many of whom are followers of Keynes and are thus referred to as **Keynesians**, regard the self-correcting mechanism, which works through wage and price adjustment, as a very slow-working mechanism because of the stickiness of wages and prices. Thus they believe it will take the economy a very long time to reach the long run, agreeing with Keynes's famous adage that "In the long-run, we are all dead." Activists therefore believe that the

government should pursue active policies to increase aggregate demand and eliminate high unemployment when it develops.

Lags and Policy Implementation

If policymakers could shift the aggregate demand curve instantaneously, activist policies could be used to move the economy immediately to the full-employment level, as we saw in the previous section. However, several types of lags prevent this immediate shift from occurring, and there are differences in the lengths of these lags for monetary versus fiscal policy.

1. The **data lag** is the time it takes for policymakers to obtain the data that describe what is happening in the economy. Accurate data on GDP, for example, are not available until several months after a given quarter is over.
2. The **recognition lag** is the time it takes for policymakers to feel confident about the signals the data are sending about the future course of the economy. For example, to minimize errors, the government will not declare the economy to be in recession until at least six months after it has determined that a recession has begun.
3. The **legislative lag** represents the time it takes to get legislation passed to implement a particular policy. The legislative lag does not apply to most monetary policy actions, such as the lowering of interest rates. It is, however, important in the implementation of fiscal policy, because it can sometimes take six months to a year to get legislation passed to change taxes or government purchases.
4. The **implementation lag** is the time it takes for policymakers to change policy instruments once they have decided on a new policy. Again, this lag is less important for the conduct of monetary policy, because the Bank of Canada can immediately change its policy interest rate, whereas it is more important for the implementation of fiscal policy. The implementation of new fiscal policy can take substantial time; for example, getting government agencies to change their spending habits takes time, as does changing tax tables.
5. The **effectiveness lag** is the time it takes for the policy to have a real impact on the economy. The effectiveness lag is both long (often a year longer) and variable (that is, substantial uncertainty exists about the length of this lag).

The existence of these lags makes the policymakers' job far more difficult and therefore weakens the case for activism. When unemployment is high, activist policy aimed at shifting the aggregate demand curve rightward to restore the economy to full employment may not produce desirable outcomes. Indeed, if the policy lags described above are very long, then by the time the aggregate demand curve shifts to the right, the self-correcting mechanism may have already returned the economy to full employment. Then, when the activist policy kicks in, it may cause output to rise above potential, leading to a rise in inflation. In situations in which policy lags are longer than the time it takes the self-correcting mechanism to work, a policy of nonactivism may produce better outcomes.

The activist/nonactivist debate came to the fore when the Obama administration advocated a fiscal stimulus package early in its administration in 2009 (see the FYI box "The Activist/Nonactivist Debate over the Obama Fiscal Stimulus Package in the United States"). In Chapter 25, we will return to the issue of just how active policy should be when we look at the role that expectations play in monetary policy.

FYI **The Activist/Nonactivist Debate over the Obama Fiscal Stimulus Package in the United States**

When President Obama took office in January 2009, he faced a very serious recession, with unemployment at over 7% and rising rapidly. Although policymakers had been using monetary policy aggressively to stabilize the economy (see Chapter 16), many activists argued that the government needed to do more by implementing a massive fiscal stimulus package. They argued that without the stimulus package, monetary policy, which had already lowered the federal funds rate to close to zero and so could not lower nominal interest rates further, would be unable to increase aggregate demand to the full-employment level. In contrast, nonactivists opposed the fiscal stimulus package, arguing that fiscal stimulus would take too long to work because of long implementation lags. They cautioned that if the fiscal stimulus kicked in after the economy had already recovered, the result would be increased volatility in inflation and economic activity.

The economics profession was split over the desirability of fiscal stimulus. Approximately 200 economists who opposed fiscal stimulus signed a petition published in the *Wall Street Journal* and the *New York Times* on January 28, 2009. An opposing petition, also signed by about 200 economists, was sent to the U.S. Congress on February 8. The Obama administration came down squarely on the side of the activists and proposed the American Recovery and Reinvestment Act of 2009, a $787 billion fiscal stimulus package that was passed by the U.S. Congress on February 13, 2009. Even after the fact, the value of the 2009 stimulus package is still hotly debated, with some believing it helped stabilize the economy and others believing it was not effective.

24.3 INFLATION: ALWAYS AND EVERYWHERE A MONETARY PHENOMENON

LO 24.3 Explain why monetary policymakers can target any inflation rate in the long run but cannot target a level of aggregate output in the long run.

Milton Friedman is famous for his adage that in the long run, "Inflation is always and everywhere a monetary phenomenon." This adage is supported by our aggregate demand and supply analysis, which asserts that monetary policymakers can target any inflation rate in the long run by shifting the aggregate demand curve through autonomous monetary policy. As an illustration, look at Figure 24-6, where the economy is at point 1, with aggregate output at potential output, so the aggregate output index is at $Y^P = 100$ and inflation is at an initial inflation target of π_1^T.

Suppose the central bank believes that this inflation target is too low and chooses to raise it to π_3^T. The central bank eases monetary policy autonomously by lowering the real interest rate at any given inflation rate, thereby increasing investment spending and aggregate demand. In Figure 24-6, the aggregate demand curve shifts to AD_3. The economy then moves to point 2, at the intersection of AD_3 and AS_1, with inflation rising to π_2. Because aggregate output is above potential output, $Y_2 > Y^P$, the short-run aggregate supply curve shifts up and to the left, eventually stopping at AS_3. The economy moves to point 3, where inflation is at the higher target level of π_3^T and the output gap is back at zero.

Mini-lecture

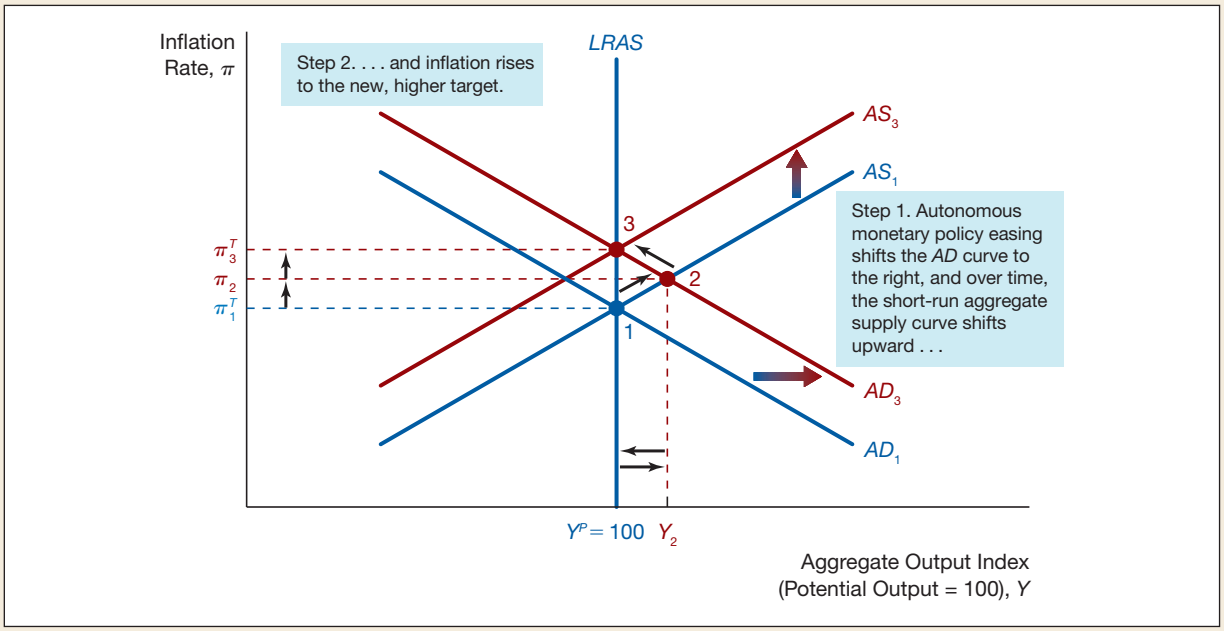


FIGURE 24-6 A Rise in the Inflation Target

To raise the inflation target to π_3^T, the central bank undertakes an autonomous monetary policy easing by lowering the real interest rate at any given inflation rate, thereby shifting the aggregate demand curve rightward to AD_3. The economy then moves to point 2, and the short-run aggregate supply curve shifts up and to the left, eventually stopping at AS_3. The economy then moves to point 3, with the output gap at zero and inflation at π_3^T.

The analysis in Figure 24-6 demonstrates the following key points:

1. ***The monetary authorities can target any inflation rate in the long run with autonomous monetary policy adjustments.***
2. ***Potential output—and therefore the quantity of aggregate output produced in the long run—is independent of monetary policy.***

24.4 CAUSES OF INFLATIONARY MONETARY POLICY

LO 24.4 Identify the sources of inflation and the role of monetary policy in propagating inflation.

If everyone agrees that high inflation is bad for an economy, why do we see so much of it? Do governments pursue inflationary monetary policies intentionally? We know that monetary authorities can set the inflation rate in the long run, so it must be that in trying to achieve other goals, governments end up with overly expansionary monetary policies and high inflation. In this section, we will examine the government policies that are the most common sources of inflation.

High Employment Targets and Inflation

One important goal of most governments is high employment, and the pursuit of this goal can bring on high inflation. For example, the U.S. government is committed by law (the Employment Act of 1946 and the Humphrey-Hawkins Act of 1978) to engage in activist policy to promote high employment. Both laws require a commitment to a high level of employment consistent with stable inflation—yet in practice, the U.S. government and the Federal Reserve have often pursued a high employment target with little concern for the inflationary consequences of such a policy. This tendency was especially apparent in the mid-1960s and 1970s, when the government and the Fed began taking an active role in attempting to stabilize unemployment.

Two types of inflation can result from an activist stabilization policy aimed at promoting high employment:

1. **Cost-push inflation** results from either a negative supply shock or a push by workers for wage hikes that are beyond those justified by productivity gains.
2. **Demand-pull inflation** results when policymakers pursue policies that increase aggregate demand.

We will now use aggregate demand and supply analysis to examine the effects of a high employment target on both types of inflation.

Cost-Push Inflation

Consider the economy shown in Figure 24-7, which is initially at point 1, the intersection of the aggregate demand curve AD_1 and the short-run aggregate supply curve AS_1. Suppose workers succeed in pushing for higher wages, either because they want to increase their real wages (their wages in terms of the goods and services they can buy) above the level justified by productivity gains or because they expect inflation to increase and therefore want their wages to increase accordingly. This cost-push shock, which acts like a negative supply shock, raises the inflation rate and shifts the short-run aggregate supply curve up and to the left, to AS_2. If the central bank takes no action to change the equilibrium interest rate and the monetary policy curve remains unchanged, the economy will move to point $2'$, at the intersection of the new short-run aggregate supply curve AS_2 and the aggregate demand curve AD_1. The aggregate output index will decline to Y', where output is below potential output, and the inflation rate will rise to $\pi_{2'}$, leading to an increase in unemployment.

In contrast, activist policymakers with a high employment target would implement policies to increase aggregate demand, such as tax cuts, increases in government purchases, or an autonomous easing of monetary policy. These policies would shift the aggregate demand curve in Figure 24-7 to AD_2, quickly returning the economy to potential output at point 2 and increasing the inflation rate to π_2. The workers would fare quite well, earning both higher wages and government protection against excessive unemployment.

The workers' success might encourage them to seek even higher wages. In addition, other workers might now realize that their wages have fallen relative to their fellow workers' wages, leading them to seek wage increases. These actions will lead to another negative supply shock, which will cause the short-run aggregate supply curve in Figure 24-7 to shift up and to the left again, to AS_3. Unemployment will develop again when we move to point $3'$, prompting additional activist policies that will shift the aggregate demand curve rightward to AD_3 and return the economy to full employment at a higher inflation rate of π_3. If this process continues, the result will be a continuing increase in inflation—a cost-push inflation.

Mini-lecture

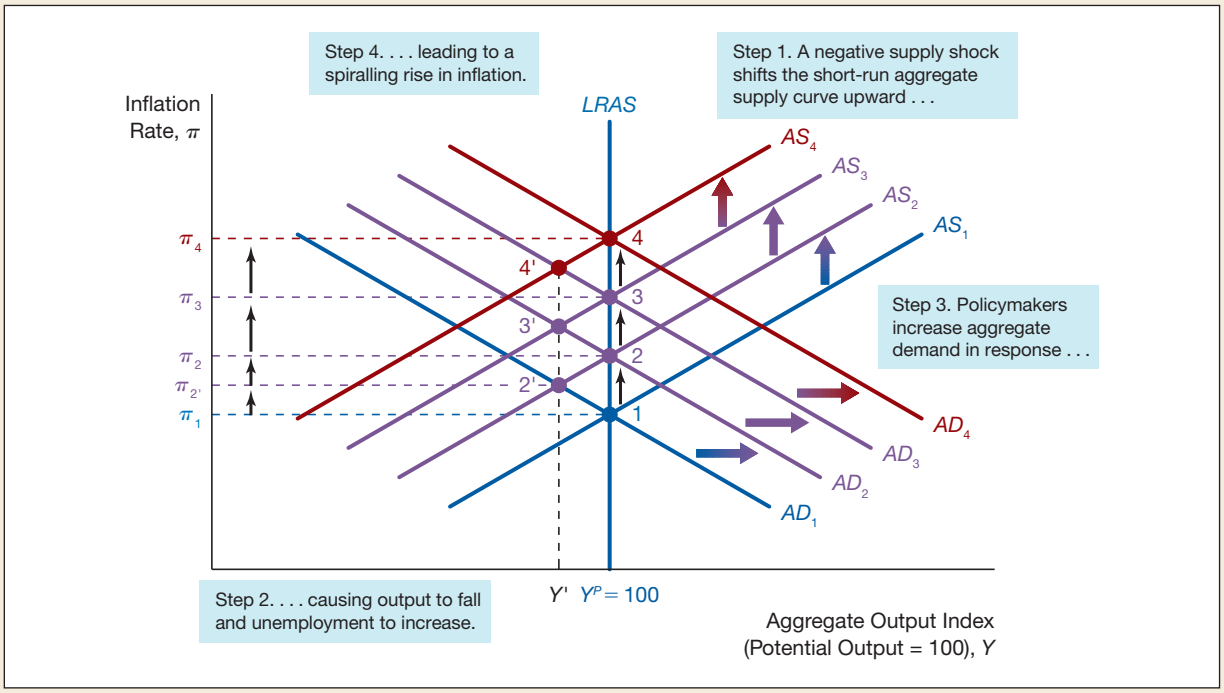


FIGURE 24-7 Cost-Push Inflation

A cost-push shock (which acts like a negative supply shock) shifts the short-run aggregate supply curve up and to the left to AS_2, and the economy moves to point 2′. To keep aggregate output at potential output and to lower the unemployment rate, policymakers shift the aggregate demand curve to AD_2 so that the economy will return quickly to potential output at point 2 and an inflation rate of π_2. Further upward and leftward shifts of the short-run aggregate supply curve, to AS_3 and beyond, lead the policymakers to continue increasing aggregate demand, leading to a continuing increase in inflation—a cost-push inflation.

Demand-Pull Inflation A goal of high employment can lead to inflationary fiscal and monetary policy in another way. Even at full employment (the natural rate of unemployment), some unemployment is always present because of frictions in the labour market that complicate the matching of unemployed workers with employers. Consequently, the unemployment rate when employment is full is still greater than zero. When policymakers mistakenly underestimate the natural rate of unemployment and so set a target for unemployment that is too low (i.e., less than the natural rate of unemployment), they set the stage for expansionary monetary policy that produces inflation.

Figure 24-8 uses aggregate supply and demand analysis to show how this scenario can unfold. If policymakers set a 4% unemployment target when the natural rate of unemployment is 5%, they are trying to achieve an output target greater than potential output. We mark this target level of the aggregate output index in Figure 24-8 as Y^T. Suppose the economy is initially at point 1: At this point, the economy is at potential output but below the target level of the aggregate output index Y^T. To hit the unemployment target of 4%, policymakers must enact policies, such as expansionary fiscal policy or an autonomous easing of monetary policy, to increase aggregate demand.

Mini-lecture

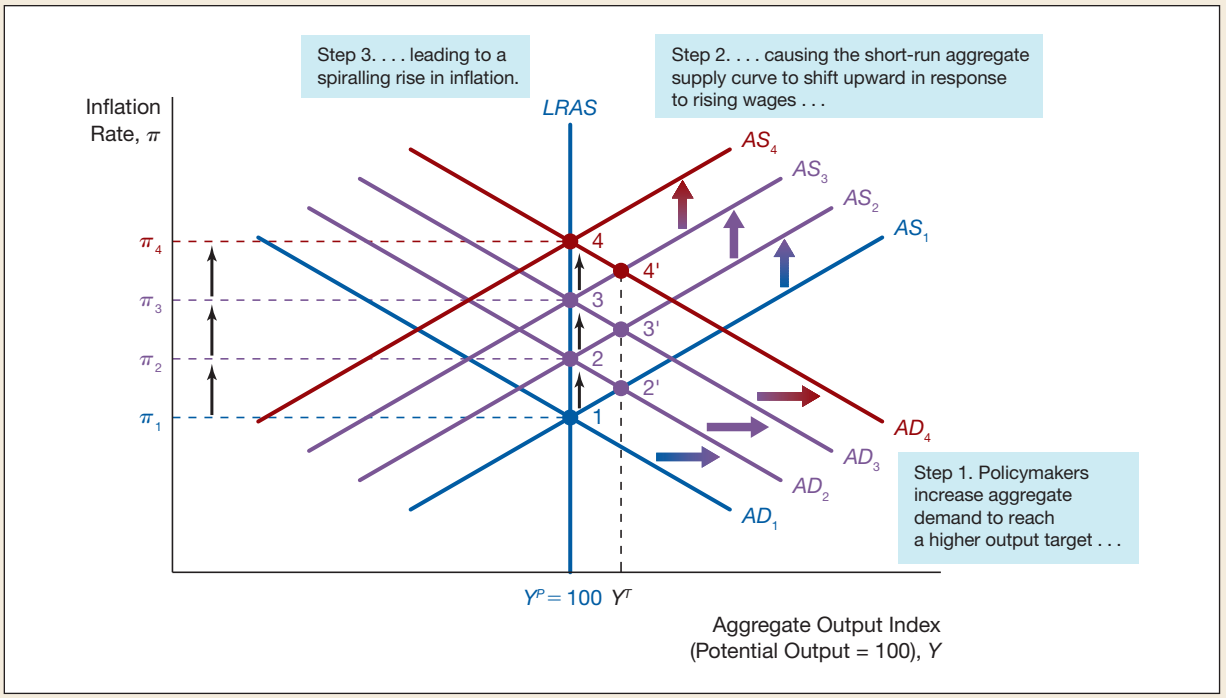


FIGURE 24-8 Demand-Pull Inflation

Too low an unemployment target (too high a target for the aggregate output index of Y^T) causes the government to increase aggregate demand, shifting the *AD* curve rightward from AD_1 to AD_2 to AD_3 and so on. Because at Y^T the unemployment rate is below the natural rate level, wages will rise and the short-run aggregate supply curve will shift up and leftward, from AS_1 to AS_2 to AS_3 and so on. The result is a continuing rise in inflation known as a demand-pull inflation.

The aggregate demand curve in Figure 24-8 shifts to the right until it reaches AD_2 and the economy moves to point 2′, where the aggregate output index is at Y^T and policymakers have achieved the 4% unemployment rate goal. But there is more to the story. At Y^T, the 4% unemployment rate is below the natural rate level and output is above potential, causing wages to rise. The short-run aggregate supply curve will now shift up and to the left, eventually reaching AS_2 and moving the economy from point 2′ to point 2, where it is back at potential output but at a higher inflation rate of π_2. Because unemployment is again higher than the target level, policymakers will once more shift the aggregate demand curve rightward to AD_3 in order to hit the output target at point 3′—and the whole process will continue to drive the economy to point 3 and beyond. The overall result is a steadily rising inflation rate.

Pursuing too low an unemployment rate target or, equivalently, too high an output target thus leads to inflationary monetary or fiscal policy. Policymakers fail on two counts: They have not achieved their unemployment target, and they have caused higher inflation. If the target rate of unemployment is below the natural rate, the process we see in Figure 24-8 will be well under way before the policymakers realize their mistake.

Cost-Push Versus Demand-Pull Inflation When inflation occurs, how do we know whether it is demand-pull inflation or cost-push inflation? We would normally expect to see demand-pull inflation when unemployment is below the natural rate level and to see cost-push inflation when unemployment is above the natural rate level. Unfortunately, economists and policymakers still struggle with accurately measuring the natural rate of unemployment. Complicating matters further, a cost-push inflation can be initiated by a demand-pull inflation, blurring the distinction. When a demand-pull inflation produces higher inflation rates, expected inflation will eventually rise and workers will demand higher wages (cost-push inflation) so that their real wages do not fall. Finally, expansionary monetary and fiscal policies produce both kinds of inflation, so we cannot distinguish between the two types of inflation on the basis of their source.

In the United States, the primary reason for inflationary policy has been policymakers' adherence to a high employment target. As we saw in Chapter 18, high inflation can also be caused by persistent government budget deficits that are financed by increases in the monetary base.

24.5 MONETARY POLICY AT THE EFFECTIVE LOWER BOUND

LO 24.5 Explain the unique challenges that monetary policymakers face at the effective lower bound, and illustrate how nonconventional monetary policy can be effective under such conditions.

So far we have assumed that a central bank can continue to lower the real interest rate as inflation falls by lowering its policy rate—say, the overnight interest rate—so that the *MP* curve is always upward-sloping. However, because the overnight interest rate is a nominal interest rate, it can never fall much below a value of zero. A negative overnight interest rate well below zero would imply that banks are willing to earn a much lower return by lending in the overnight funds market than they could earn by holding cash in their vaults, with its zero rate of return. As discussed in Chapter 16, the floor on the policy rate is referred to as the *effective lower bound*, and it creates a particular problem for the conduct of monetary policy.

Deriving the Aggregate Demand Curve with the Effective Lower Bound

To understand more deeply the problems created by the effective lower bound with regard to the conduct of monetary policy, let's look at what happens to the aggregate demand curve when a central bank cannot lower its policy rate much below zero. A key fact in this analysis is that the real interest rate is the interest rate that is adjusted for expected inflation, that is, the real rate equals the nominal rate minus expected inflation, $r = i - \pi^e$. Panel (a) of Figure 24-9 shows an *MP* curve that hits the effective lower bound, which, for simplicity, we will assume is zero, although it could be slightly less than zero. For our purposes, let's assume that expected inflation moves closely with the actual inflation rate (as it usually does), as indicated on the horizontal axis in panel (a). Let's start at point 3 on the *MP* curve, where inflation is at 3% and the real interest rate is at 2%. Now let's see what happens as inflation falls from 3% to 2%. At this inflation level, the monetary authorities will

Mini-lecture

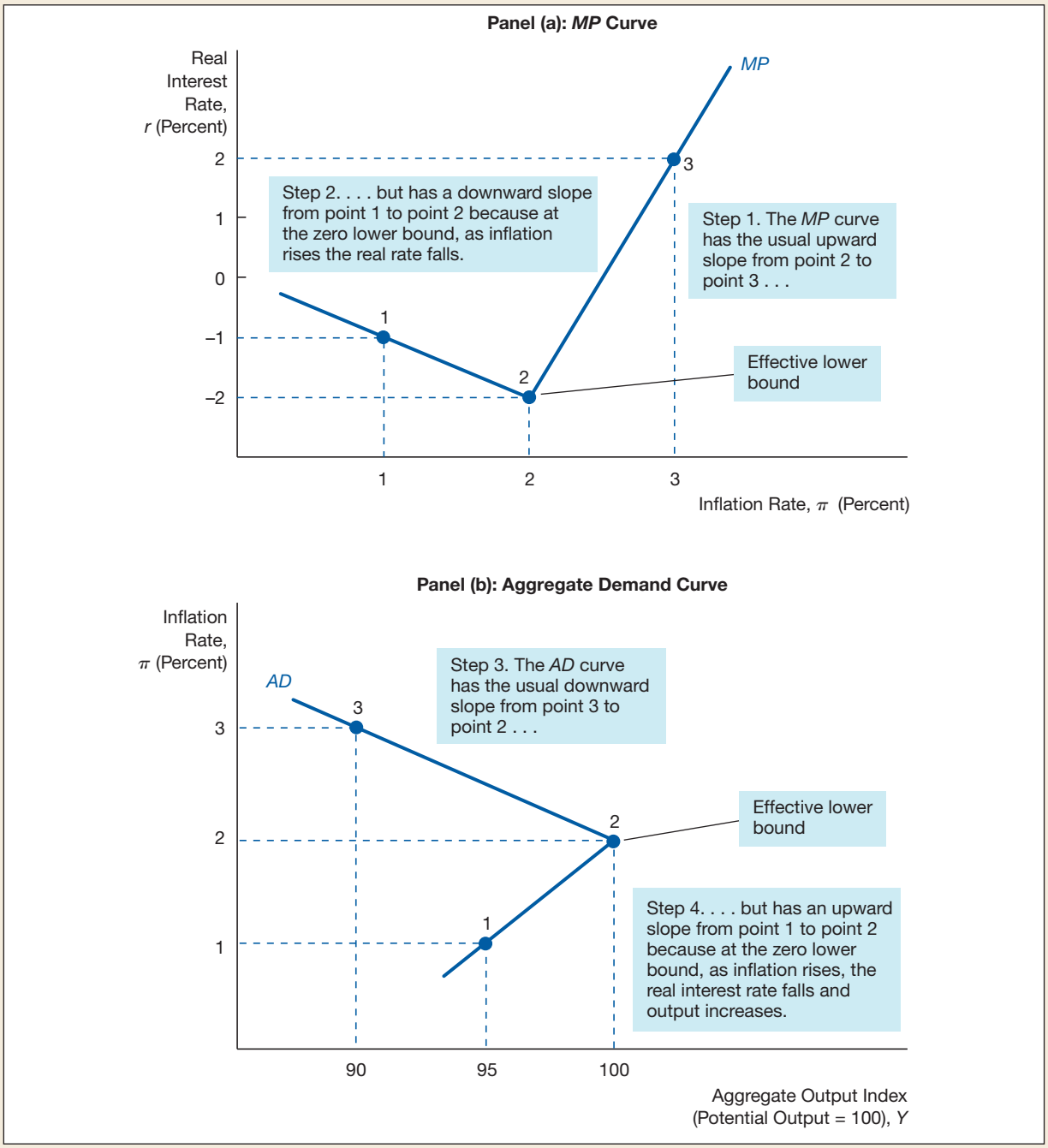

Panel (a): MP Curve
Real Interest Rate, r (Percent)
MP
2
1
0
–1
–2
3
1
2
Step 2. . . . but has a downward slope from point 1 to point 2 because at the zero lower bound, as inflation rises the real rate falls.
Step 1. The MP curve has the usual upward slope from point 2 to point 3 . . .
Effective lower bound
1
2
3
Inflation Rate, π (Percent)
Panel (b): Aggregate Demand Curve
Inflation Rate, π (Percent)
AD
3
2
1
Step 3. The AD curve has the usual downward slope from point 3 to point 2 . . .
Effective lower bound
Step 4. . . . but has an upward slope from point 1 to point 2 because at the zero lower bound, as inflation rises, the real interest rate falls and output increases.
90
95
100
Aggregate Output Index (Potential Output = 100), Y

FIGURE 24-9 Derivation of the Aggregate Demand Curve with an Effective Lower Bound

In panel (a), the *MP* curve has the usual upward slope in going from point 2 to point 3, but it slopes downward in the segment from point 1 to point 2 because, with the policy rate at the floor of the effective lower bound, as inflation and expected inflation fall, the real interest rate rises. This produces a kink in the aggregate demand curve, as shown in panel (b).

want to lower the real interest rate to, say, −2% (point 2 in panel (a)), which will require them to lower the policy rate to zero $(r = 0 - 2\% = -2\%)$. At point 2 in panel (a), the effective lower bound has been reached on the *MP* curve. So far, our analysis is identical to the analysis given in Chapter 22, with the MP curve having the usual upward slope.

Now, what happens if the inflation rate falls even further—say, to 1%? The monetary authorities would like to lower the real interest rate by lowering the policy rate again, but they can't because the policy rate has already hit the floor of the effective lower bound (which we are assuming is equal to zero). Indeed, as indicated by point 1 on the *MP* curve, with the nominal policy rate at the effective lower bound at zero, the real interest rate at an inflation rate of 1% has now risen to −1% $(r = 0 - 1\% = -1\%)$. Hence, we see that the segment of the *MP* curve that connects point 2 to point 1 is downward-sloping, the opposite of what we found in Chapter 22.

Now let's see what happens to the aggregate demand curve, shown in panel (b) of Figure 24-9. At a 3% inflation rate and a real interest rate of 2% (point 3 on the *MP* curve in panel (a)), the equilibrium level of the aggregate output index is, say, 90, marked as point 3 on the aggregate demand curve in panel (b). Now, when inflation falls to 2% and the real interest rate is at −2%, as indicated by point 2 on the *MP* curve in panel (a), the aggregate output index rises to 100 because planned investment spending rises with the lower real interest rate. The inflation rate of 2% and level of the aggregate output index is marked as point 2 on the *AD* curve in panel (b). This point is also at the effective lower bound. The *AD* curve has the usual downward slope from point 3 to point 2.

However, if inflation falls to 1%, point 1 on the *MP* curve indicates that the real interest rate will rise to −1% because, at the effective lower bound, where the nominal policy rate is stuck near zero, the decrease in both inflation and expected inflation will raise the real interest rate. The rise in the real interest rate will cause the aggregate output index to fall, to 95 at point 1 on the *AD* curve in panel (b). Therefore, in going from point 1 to point 2, the aggregate demand curve slopes upward rather than downward. The presence of the effective lower bound thus can produce a kinked aggregate demand curve of the type seen in panel (b).

The Disappearance of the Self-Correcting Mechanism at the Effective Lower Bound

Now let's analyze what happens in our aggregate demand and supply diagram when the economy is hit by a large negative shock, such as the shock that occurred during the global financial crisis from 2007 to 2009, so that the effective lower bound becomes binding. In this situation, the initial short-run aggregate supply curve intersects the upward-sloping part of the aggregate demand curve at point 1 in Figure 24-10, where the aggregate output index is at Y_1, which is less than $Y^P = 100$. Because $Y_1 < Y^P$, there is slack in the economy, and so the short-run aggregate supply curve will fall to AS_2 and the economy will move to point 2, at the intersection of the AS_2 and *AD* curves, where inflation and the aggregate output index will have declined to π_2 and Y_2, respectively. Now, Y_2 is even lower relative to Y^P, so the short-run aggregate supply curve will shift down even further, to AS_3, and the economy will move to point 3, where inflation and the aggregate output index have fallen even further, to π_3 and Y_3, respectively.

Mini-lecture

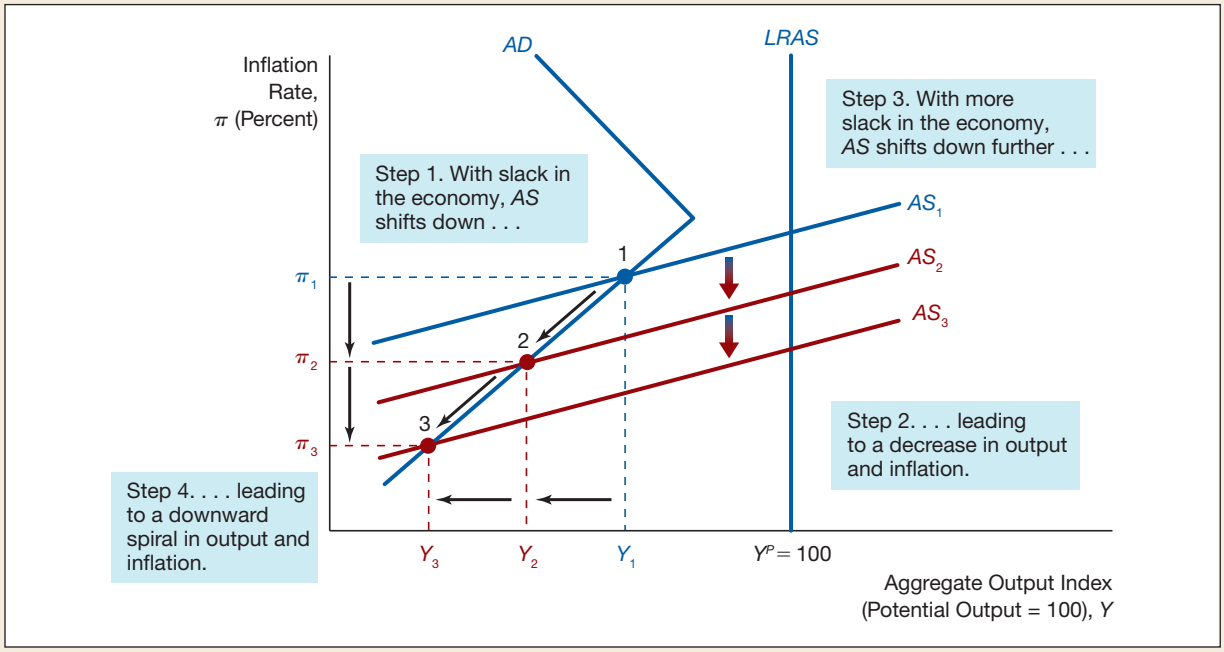

AD
LRAS
Inflation Rate, π (Percent)
Step 1. With slack in the economy, AS shifts down . . .
Step 3. With more slack in the economy, AS shifts down further . . .
AS1
AS2
AS3
π1
π2
π3
1
2
3
Step 2. . . . leading to a decrease in output and inflation.
Step 4. . . . leading to a downward spiral in output and inflation.
Y3
Y2
Y1
YP = 100
Aggregate Output Index (Potential Output = 100), Y

FIGURE 24-10 The Absence of the Self-Correcting Mechanism at the Effective Lower Bound

At the initial equilibrium at point 1, $Y_1 < Y^P$, so the short-run aggregate supply curve shifts down to AS_2 and the economy moves to point 2, where the aggregate output index and inflation have fallen to Y_2 and π_2, respectively. Because Y_2 is even lower relative to Y^P, the short-run aggregate supply curve shifts down even further to AS_3, and the economy moves to point 3, where the aggregate output index and inflation have fallen even further, to Y_3 and π_3, respectively. Both output and inflation therefore experience downward spirals.

Our analysis of Figure 24-10 reveals two key results:

- First, the self-correcting mechanism is no longer operational. When the economy is in a situation in which equilibrium output is below potential output and the effective lower bound on the policy rate has been reached, output is not restored to its potential level if policymakers do nothing. Indeed, the opposite occurs—the economy goes into a downward spiral.
- Second, in this situation, the economy goes into a deflationary spiral, characterized by continually falling inflation and output.

The intuition behind these two results is fairly straightforward. When output is below potential and the policy rate has hit the floor of the effective lower bound, the resulting fall in inflation leads to higher real interest rates, which depress output further, which causes inflation to fall further, and so on. Schematically, this chain of events can be expressed as follows:

$$Y < Y^P \Rightarrow \pi\downarrow \Rightarrow r\uparrow \Rightarrow Y\downarrow \Rightarrow Y \ll Y^P \Rightarrow \pi\downarrow \Rightarrow r\uparrow \Rightarrow Y\downarrow$$

The final outcome is that both output and inflation go into downward spirals.

APPLICATION

Nonconventional Monetary Policy and Quantitative Easing

At the effective lower bound, conventional expansionary monetary policy is no longer an option because the monetary policy authorities are unable to lower the policy rate. As a result, the central bank has to turn to the nonconventional policies discussed in Chapter 16 to stimulate the economy. Here we will analyze how these different nonconventional policies lead to economic expansion while avoiding the downward spirals in output and inflation that characterized the situation described in Figure 24-10.

Recall that nonconventional monetary policy takes three forms: liquidity provision, asset purchases (quantitative easing), and management of expectations. To see how each of these forms works, recall from Chapter 21 that the real interest rate for investments r_i reflects not only the short-term real interest rate set by the central bank, r, but also an additional term $\bar{f}$, which we referred to previously as financial frictions. This relationship can be written mathematically as follows

$$r_i = r + \bar{f} \tag{1}$$

Each of these nonconventional monetary policy measures helps raise aggregate output and inflation by lowering $\bar{f}$ in the *AD/AS* model. Let's look at each of these measures in turn.

Liquidity Provision

The effective lower bound situation depicted in Figure 24-11 often arises when credit markets seize up and there is a sudden shortage of liquidity, as occurred during the 2007–2009 global financial crisis. The shortage of liquidity results in a sharp rise in financial frictions, which shifts the aggregate demand curve to AD_1 in Figure 24-11 to point 1, where it intersects the aggregate supply curve and where the policy rate has hit a floor of the effective lower bound and output is below potential. A central bank can bring down financial frictions directly by increasing its lending facilities in order to provide more liquidity to impaired markets so that they can return to their normal functions, thereby bringing down the $\bar{f}$ term. As we saw in Chapter 23, this decline in financial frictions lowers the real interest rate for investments, $r_i = r + \bar{f}$, and so increases investment spending and the quantity of aggregate output demanded at any given inflation rate. The aggregate demand curve then shifts to the right to AD_2 and the economy moves to point 2, where both output and inflation have risen. Indeed, if liquidity provisions are sufficiently successful, the economy can move back to its full-employment level, where output will return to potential, as indicated by point 2 in the figure.

Asset Purchases and Quantitative Easing

The monetary authorities can also lower the $\bar{f}$ term by lowering credit spreads through the purchase of privately issued securities. When the monetary authorities purchase a private security, such a purchase raises the security's price and therefore lowers its interest rate, thereby lowering the credit spread and hence $\bar{f}$ and the real interest rate for investments. The decline in the real interest rate for investments at any given inflation rate then causes the aggregate demand curve to shift to the right, as shown in Figure 24-11, and raises both output and inflation.

Mini-lecture

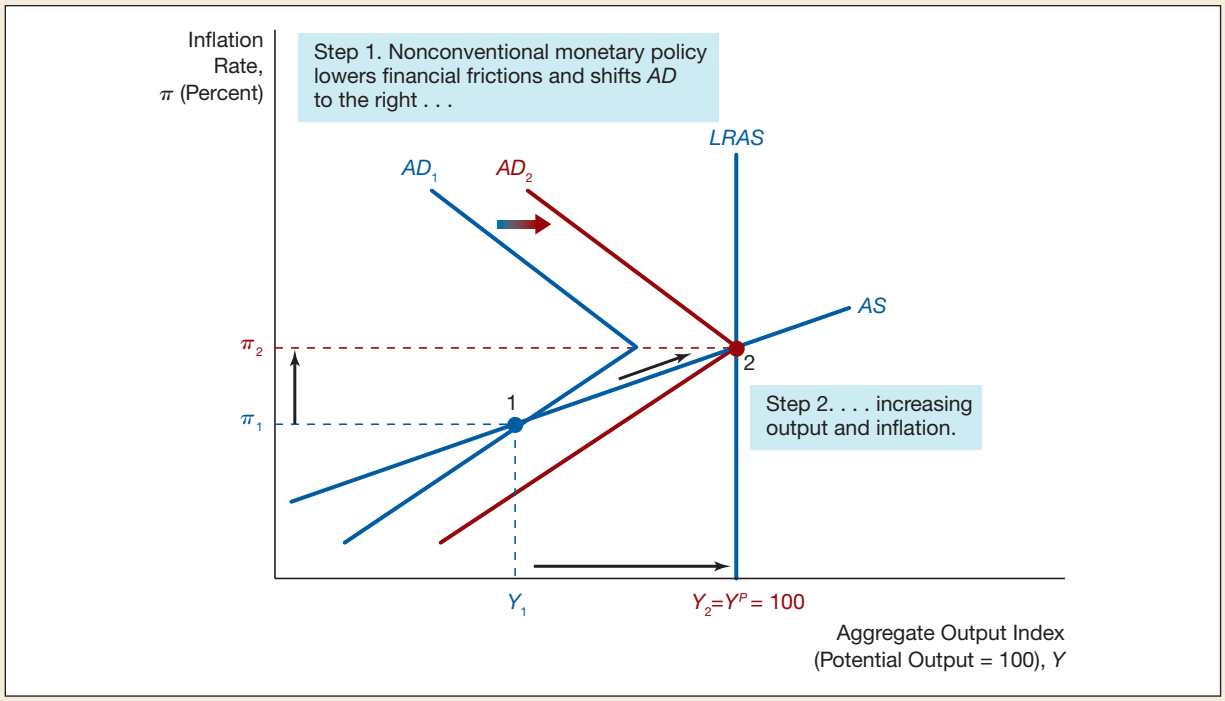


FIGURE 24-11 Response to Nonconventional Monetary Policy

Nonconventional monetary policy, whether it involves liquidity provision, asset purchases, or management of expectations, lowers $\bar{f}$, which in turn lowers the real interest rate for investments at any given inflation rate and shifts the aggregate demand curve to AD_2. The economy moves to point 2, where the aggregate output index and inflation have risen to Y_2 and π_2, respectively.

Because investments typically are associated with long-term projects, the real interest rate for investments is likely to be a long-term interest rate and therefore differs from the short-term real interest rate r. Hence the $\bar{f}$ term in Equation 1 can be viewed as reflecting not only financial frictions and credit spreads but also the spread between long-term and short-term rates. This means that asset purchases of long-term government securities also can lower the real interest rate for investments. When the Bank of Canada purchases long-term Canada bonds, for example, this action raises their price and lowers long-term interest rates. The result is a decline in $\bar{f}$ and the real interest rate for investments at any given inflation rate, and so the aggregate demand curve will shift to the right, to AD_2 in Figure 24-11, thereby raising output and inflation.

When a central bank engages in liquidity provision or asset purchases, its balance sheet necessarily expands. Indeed, as we saw in Chapter 16, from before the financial crisis began in September 2007 through 2014, the value of Federal Reserve assets rose from about US$800 billion to over US$4 trillion and then rose again dramatically starting in March 2020 in response to the coronavirus pandemic. In Canada, before the Covid-19 pandemic, the Bank of Canada owned about $100 billion worth of Canada bonds, which was about 15% of the total outstanding. In response to the Covid-19 crisis, on April 1, 2020, the Bank of Canada started buying at least $5 billion per week of Canada bonds and pledged to continue that pace of purchases until the economy recovers. Such an expansion

of the central bank's balance sheet is referred to as *quantitative easing* because it leads to a huge increase in liquidity in the economy, which can be a powerful force in stimulating the economy in the near term and possibly producing inflation down the road.

However, an expansion in the central bank's balance sheet in and of itself may not be enough to stimulate the economy. As we saw in our *AD/AS* analysis of the effective lower bound, unless quantitative easing is able to lower the real interest rate for investments, in what former Fed chair Ben Bernanke referred to as a *credit easing*, there will be no impact on the aggregate demand curve and hence no impact on output and inflation. If the asset-purchase program involves only the purchase of short-term government securities, the program is unlikely to affect credit spreads or the spread between long- and short-term interest rates, and so $\bar{f}$ and the real interest rate for investments will remain unchanged. The result will be a minimal impact on the aggregate economy.[1] Indeed, this was the experience in Japan when the Bank of Japan pursued a large-scale asset-purchase program that primarily involved purchases of short-term government bonds. Not only did the economy fail to recover but inflation even turned negative.

Management of Expectations

One form of management of expectations is forward guidance in which the central bank commits to keeping the policy rate low for a long period of time, another way of lowering long-term interest rates relative to short-term rates and thereby lowering $\bar{f}$ and the real interest rate for investments. Because investors can choose to invest in a long-term bond rather than investing in a sequence of short-term bonds, as we saw in Chapter 6, the interest rate on long-term bonds will be closely related to an average of the short-term interest rates that markets expect to occur over the life of the long-term bond. By committing to the future policy action of keeping the overnight interest rate at zero for an extended period, a central bank can lower the market's expectations of future short-term interest rates, thereby causing the long-term interest rate to fall. The result will be a decline in $\bar{f}$ and the real interest rate for investments, which will shift the aggregate demand curve to the right as in Figure 24-11, raising both output and inflation.

So far, the mechanisms for the efficacy of nonconventional monetary policies have operated through the $\bar{f}$ term and rightward shifts in the aggregate demand curve as in Figure 24-11. However, management of expectations can also operate through shifts in the short-run aggregate supply curve that can be achieved by raising expectations of inflation, as shown in Figure 24-12. Recall from Chapter 23 that a rise in inflation expectations—say, because the central bank commits to doing whatever it takes to raise inflation in the future—will shift the short-run aggregate supply curve up, to AS_2 in Figure 24-12, moving the economy to point 2, where the aggregate output index and inflation rise to Y_2 and π_2, respectively. The intuition behind this result is straightforward: With the policy rate fixed at the effective lower bound, the rise in expected inflation will lead to a decline in the real interest rate, which will cause investment spending

[1]There are two other reasons that quantitative easing in and of itself will not necessarily be stimulative. First, large expansions in a central bank's balance sheet do not necessarily result in large increases in the money supply. As indicated in the application in Chapter 15, "Quantitative Easing and the Money Supply, 2007–2020," this was exactly the case in the United States from 2007 to 2014 and again in 2020, when huge expansions in the Fed's balance sheet and the monetary base did not result in a large increase in the money supply because most of the increase in the monetary base just flowed into holdings of excess reserves. Second, an increase in the monetary base does not mean that banks will increase lending, because they have the option of just adding to their holdings of excess reserves instead of making loans. This is exactly what happened during the global financial crisis and the coronavirus pandemic, when a huge increase in the monetary base led primarily to a massive rise in excess reserves, and bank lending did not increase.

Mini-lecture

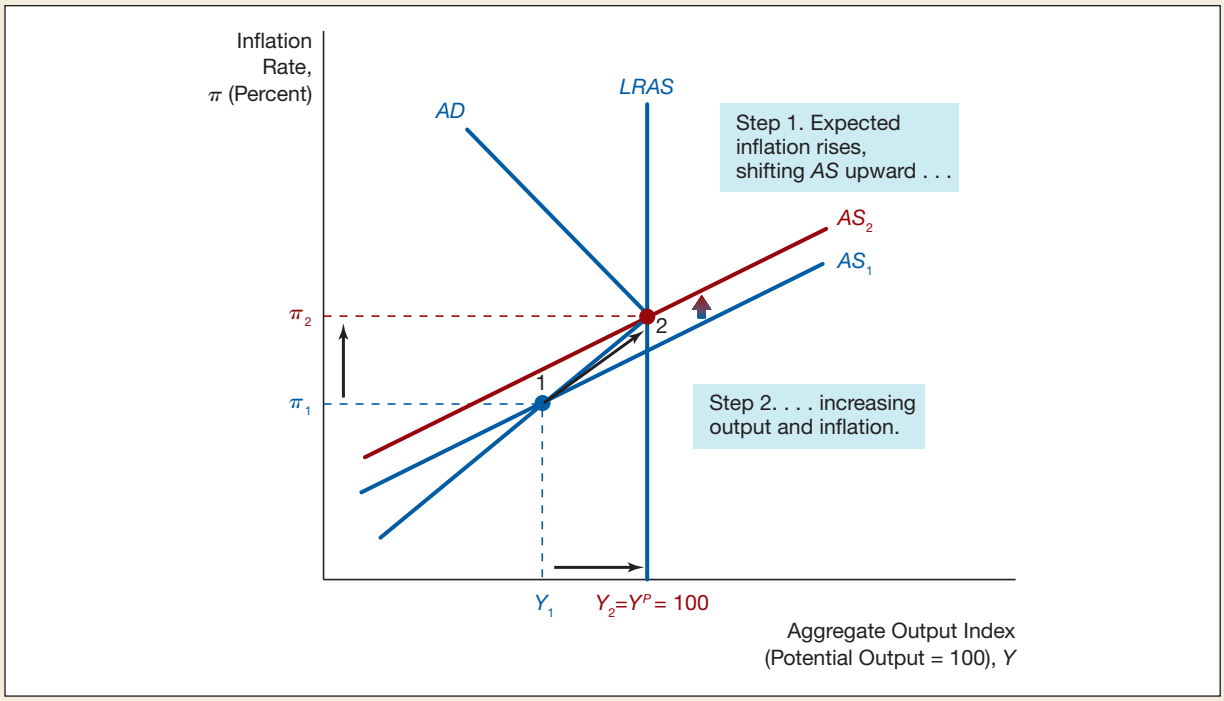


FIGURE 24-12 Response to a Rise in Inflation Expectations

A rise in inflation expectations causes the short-run aggregate supply curve to shift up to AS_2 and the economy to move to point 2, where the aggregate output index and inflation rise to Y_2 and π_2, respectively.

and aggregate output to rise as the economy slides up the aggregate demand curve from point 1 to point 2, as shown in the figure. One problem with this strategy, however, is that the public must believe that inflation will actually rise in the future. If the central bank's commitment to raising inflation is not credible, then inflation expectations may not rise and this particular type of management of expectations will not work.

APPLICATION Abenomics and the Shift in Japanese Monetary Policy in 2013

By 2012, the Japanese economy had been in a funk for well over 10 years, with very low growth, the policy rate stuck at the effective lower bound near zero, and the economy experiencing deflation. With this scenario as his backdrop, Shinzo Abe won the general election for prime minister of Japan in December 2012, and after taking office he promoted a major shift in economic policy aimed at stimulating economic growth, a policy that the media has since dubbed "Abenomics." A key element of Abenomics was a sea change in monetary policy. First, Abe pressured the Bank of Japan to double its inflation target from 1% to 2% in January 2013, over the objections of the former governor of the Bank of Japan, Masaaki Shirakawa, who then resigned in March. After taking over the former governor's position in March 2013, the new Bank of Japan governor,

Mini-lecture

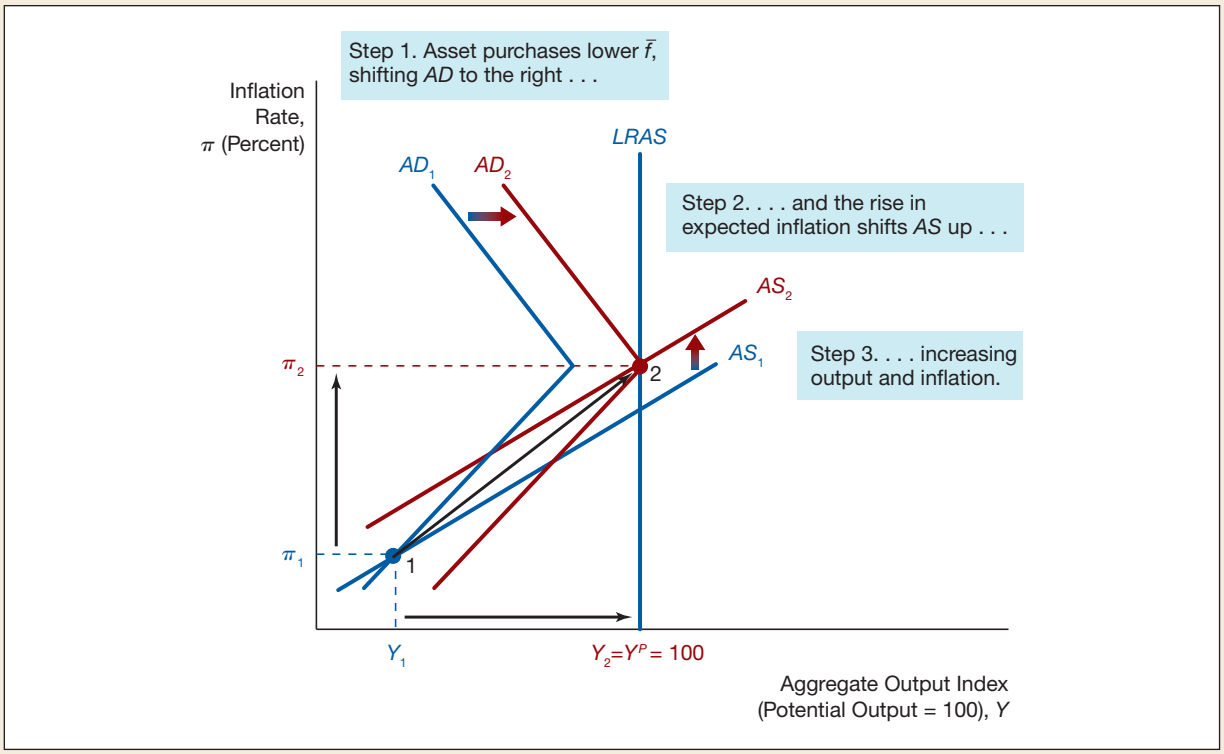


FIGURE 24-13 Response to the Shift in Japanese Monetary Policy in 2013

The Bank of Japan's revised asset-purchase program lowers $\bar{f}$, which lowers the real interest rate on investments at any given inflation rate and shifts the aggregate demand curve to the right, from AD_1 to AD_2. The rise in inflation expectations causes the short-run aggregate supply curve to shift upward from AS_1 to AS_2. The economy moves to point 2, where the aggregate output index and inflation rise to Y_2 and π_2, respectively.

Haruhiko Kuroda, announced a major change in the way the Bank of Japan would conduct monetary policy in the future. First, in contrast to the previous governor, who never formally committed to achieving the 1% inflation target, Kuroda committed to achieving the higher 2% inflation objective within two years. Second, he indicated that the Bank of Japan would now engage in a massive asset-purchase (quantitative easing) program that would not only double the size of the Bank of Japan's balance sheet but would also involve the purchase of a very different set of assets. Specifically, rather than purchasing short-term government bonds, the Bank of Japan would now purchase long-term bonds, including private securities such as real estate investment trusts.

We can use our analysis from the previous section to predict why this approach was taken to fix the Japanese economy. First, in contrast to the previous quantitative easing, the Abenomics program sought to lower $\bar{f}$ through the purchase of long-term assets. Specifically, the program would lower $\bar{f}$ by lowering credit spreads through the purchase of private securities and also by lowering long-term interest rates through the purchase of long-term government bonds. As we have seen, with the policy rate at the effective lower bound, a lower $\bar{f}$ would lead to a lower real interest rate for investments, thereby shifting the aggregate demand curve to the right, to AD_2 in Figure 24-13.

Second, the higher inflation target, and even more importantly the stronger commitment by Kuroda to achieving this higher target, should, according to our analysis, raise expected inflation and hence shift the short-run aggregate supply curve to AS_2. As we can see in Figure 24-13, the economy would then move to point 2, where both output and inflation would rise.

In other words, the new monetary policy's two-pronged attack would both lower the real interest rate for investments directly through the asset-purchase program and also directly raise inflation expectations, providing another factor that would drive down the real interest rate. Both of these mechanisms would then operate to promote an economic expansion and an exit from the deflationary environment that the Japanese had been experiencing for the past 15 years. Although this new policy strategy has helped raise Japanese inflation above zero and has led to lower unemployment, it has not been successful in raising Japanese inflation to the 2% target. This failure illustrates how difficult it is to raise inflation expectations once a prolonged period of deflation has occurred. ◆

SUMMARY

1. For aggregate demand shocks, the price stability and economic activity stability objectives are consistent: Stabilizing inflation stabilizes economic activity, even in the short run. For supply shocks, however, there is a trade-off between stabilizing inflation and stabilizing economic activity in the short run. In the long run, however, there is no conflict between stabilizing inflation and stabilizing economic activity.
2. Activists regard the self-correcting mechanism, which works through wage and price adjustment, as very slow and hence feel that the government should pursue active, accommodating policy to address high unemployment when it develops. Nonactivists, by contrast, believe that the self-correcting mechanism works quickly and therefore advocate that the government should avoid the implementation of active policies aimed at eliminating unemployment.
3. Milton Friedman's view that in the long run inflation is always and everywhere a monetary phenomenon is borne out by aggregate demand and supply analysis: Such analysis shows that monetary policymakers can target any inflation rate in the long run through autonomous monetary policy, which makes use of the overnight interest rate policy tool to change the level of aggregate demand and thereby adjust the equilibrium inflation rate.
4. Two types of inflation can result from an activist stabilization policy aimed at promoting high employment: cost-push inflation, which is caused by negative supply shocks or a push by workers for wages that are higher than the level justified by productivity gains; and demand-pull inflation, which results when policymakers pursue high output and employment targets through policies that increase aggregate demand.
5. When the nominal policy rate hits the floor of the effective lower bound, the aggregate demand curve becomes upward-sloping, which means that the self-correcting mechanism that returns the economy to full employment is no longer operational. At the effective lower bound, in order to boost output and inflation, the monetary authorities must turn to nonconventional policies of three types: liquidity provision, asset purchases (typically referred to as quantitative easing), and management of expectations.

MyLab Economics Make the grade with MyLab Economics: Select end-of-chapter questions can be found on MyLab Economics. You can practise them as often as you want, and most feature step-by-step guided instructions to help you find the right answer.

QUESTIONS

1. What does it mean when we say that the inflation gap is negative?
2. "If autonomous spending falls, the central bank should lower its inflation target in order to stabilize inflation." Is this statement true, false, or uncertain? Explain your answer.
3. For each of the following shocks, describe how monetary policymakers would respond (if at all) to stabilize economic activity. Assume the economy starts at a long-run equilibrium.
   a. Consumers reduce autonomous consumption.
   b. Financial frictions decrease.
   c. Government spending increases.
   d. Taxes increase.
   e. The domestic currency appreciates.
4. During the global financial crisis, how was the Bank of Canada able to help offset the sharp increase in financial frictions without the option of lowering interest rates further? Did the Bank's plan work?
5. Why does the divine coincidence simplify the job of policymakers?
6. Why do negative supply shocks pose a dilemma for policymakers?
7. Suppose three economies are hit with the same negative supply shock. In country A, inflation initially rises and output falls; then inflation rises more and output increases. In country B, inflation initially rises and output falls; then both inflation and output fall. In country C, inflation initially rises and output falls; then inflation falls and output eventually increases. What type of stabilization approach did each country take?
8. "Policymakers would never respond by stabilizing output in response to a positive supply shock." Is this statement true, false, or uncertain? Explain your answer.
9. The fact that it takes a long time for firms to get new plants and equipment up and running is an illustration of what policy problem?
10. In the United States, many observers have commented in recent years on the "political gridlock in Washington, D.C.," and referred to Congress as a "Do Nothing Congress." What type of policy lag is this describing?
11. Is stabilization policy more likely to be conducted through monetary policy or through fiscal policy? Why?
12. "If the data and recognition lags could be reduced, activist policy probably would be more beneficial to the economy." Is this statement true, false, or uncertain? Explain your answer.
13. If the economy's self-correcting mechanism works slowly, should the government necessarily pursue discretionary policy to eliminate unemployment? Why or why not?
14. In the early 1970s, an unexpected productivity slowdown led most economists to believe that potential output was larger than it actually was. How are policymakers likely to respond under this situation, and what do you expect to be the outcome?
15. In both the 1960s and the 1970s in Canada, inflation increased significantly. However, unemployment behaved very differently in the two decades. Why is this?
16. Given a relatively steep and a relatively flat short-run aggregate supply curve, which curve would support the case for nonactivist policy? Why?
17. "Because government policymakers do not consider inflation desirable, their policies cannot be the source of inflation." Is this statement true, false, or uncertain? Explain your answer.
18. How can monetary authorities target any inflation rate they wish?
19. What will happen if policymakers erroneously believe that the natural rate of unemployment is 7% when it is actually 5% and therefore pursue stabilization policy?
20. How can demand-pull inflation lead to cost-push inflation?
21. How does the policy rate hitting a floor a little below zero lead to an upward-sloping aggregate demand curve?
22. Why does the self-correcting mechanism stop working when the policy rate hits the effective lower bound?
23. In what ways can nonconventional monetary policy affect the real interest rate for investments when the economy reaches the effective lower bound? How are credit spreads affected?

APPLIED PROBLEMS

24. Suppose the government decides to decrease government expenditures as a means of cutting the existing government budget deficit.

a. Using a graph of aggregate demand and supply, show the effects of such a decision on the economy in the short run. Describe the effects on inflation and output.

b. What will be the effect on the real interest rate, the inflation rate, and the aggregate output index if the Bank of Canada decides to stabilize the inflation rate?

25. Use a graph of aggregate demand and supply to demonstrate how lags in the policy process can result in undesirable fluctuations in output and inflation.

26. As monetary policymakers become more concerned with inflation stabilization, the slope of the aggregate demand curve becomes flatter. How does the resulting change in the slope of the aggregate demand curve help stabilize inflation when the economy is hit with a negative supply shock? How does this affect output? Use a graph of aggregate demand and supply to demonstrate.

27. In 2003, as the U.S. economy finally seemed poised to exit its ongoing recession, the Fed began to worry about a "soft patch" in the economy, in particular the possibility of a deflation. As a result, the Fed proactively lowered the federal funds rate from 1.75% in late 2002 to 1% by mid-2003, the lowest federal funds rate on record up to that point in time. In addition, the Fed committed to keeping the federal funds rate at this level for a considerable period of time. This policy was considered highly expansionary and was seen by some as potentially inflationary and unnecessary.

a. How might fears of the effective lower bound justify such a policy, even if the economy was not actually in a recession?

b. Show the impact of these policies on the *MP* curve and the *AD/AS* graph. Be sure to show the initial conditions in 2003 and the impact of the policy on the deflation threat.

28. Suppose that $\bar{f}$ is determined by two factors: financial panic and asset purchases.

a. Using an *MP* curve and an *AS/AD* graph, show how a sufficiently large financial panic can pull the economy below the effective lower bound and into a destabilizing deflationary spiral.

b. Using an *MP* curve and an *AS/AD* graph, show how a sufficient amount of asset purchases can reverse the effects of the financial panic depicted in part (a).

DATA ANALYSIS PROBLEMS

1. **Real-Time Data Analysis** Go to the St. Louis Federal Reserve FRED database and find quarterly data on the Canadian unemployment rate (aged 15 and over) and the quarterly Canadian consumer price index (all items) since 1961Q1 for both series.

a. For the most recent four quarters of data available, calculate the average inflation gap using the Bank of Canada's 2% inflation target.

b. For the most recent four quarters of data available, calculate the average unemployment gap, using 5.6% as the presumed natural rate of unemployment.

c. Based on your answers to parts (a) and (b), does the divine coincidence apply to the current economic situation in Canada? Why or why not? What does your answer imply about the sources of shocks that have impacted the current economy? Briefly explain.

2. **Real-Time Data Analysis** Go to the St. Louis Federal Reserve FRED database and find U.S. data on the personal consumption expenditure price index (PCECTPI), the unemployment rate (UNRATE), and an estimate of the natural rate of unemployment (NROU). For the price index, adjust the *units* setting to "Percent Change From Year Ago." For the unemployment rate, adjust the *frequency* setting to "Quarterly." Select the data from 2000 through the most current data available, download the data, and plot all three variables on the same graph. Using your graph, identify periods of demand-pull or cost-push movements in the inflation rate. Briefly explain your reasoning.

25 The Role of Expectations in Monetary Policy

Learning Objectives

25.1 Summarize the Lucas critique.

25.2 Compare and contrast the use of policy rules versus discretionary policy.

25.3 Summarize and illustrate the benefits of a credible central bank.

25.4 Identify the ways in which central banks can establish and maintain credibility.

Preview

After World War II, economists, armed with models like the ones we developed in Chapters 21–24, felt that discretionary policies could reduce the severity of business cycle fluctuations without creating inflation. In the 1960s and 1970s, these economists got their chance to put their policies into practice, but the results were not what they had anticipated. The economic record for that period is not a happy one: Inflation accelerated, with the inflation rate often climbing to above 10%, and unemployment figures deteriorated from those of the 1950s.[1]

In the 1970s and 1980s, economists, including Nobel Prize winners Robert Lucas of the University of Chicago and Thomas Sargent, now at New York University, used the rational expectations theory discussed in Chapter 7 to examine why discretionary policies appear to have performed so poorly. Their analysis cast doubt on whether macroeconomic models can be used to evaluate the potential effects of policy and on whether policy can be effective when the public *expects* that it will be implemented. Because the analysis of Lucas and Sargent has such strong implications for the way in which policy should be conducted, it has been labelled the *rational expectations revolution*.[2]

This chapter examines the analysis behind the rational expectations revolution. We start first with the Lucas critique, which indicates that because expectations are an important part of economic behaviour, it may be quite difficult to predict what the outcome of a discretionary policy will be. We then discuss the effect of rational expectations on the aggregate demand and supply analysis we developed in Chapter 23 and explore how this theoretical breakthrough has shaped current policymaking models and debates.

25.1 LUCAS CRITIQUE OF POLICY EVALUATION

LO 25.1 Summarize the Lucas critique.

Economists have long used **macroeconometric models** to forecast economic activity and to evaluate the potential effects of policy options. In essence, the models are collections of equations that describe statistical relationships among many economic variables. Economists can feed data into such models, which then churn out a forecast or prediction.

[1]Some of the deterioration can be attributed to supply shocks in 1973–1975 and 1978–1980.

[2]Other economists who have been active in promoting the rational expectations revolution are Robert Barro of Harvard University, Bennett McCallum of Carnegie-Mellon University, Nobel Prize winner Edward Prescott of Arizona State University, and Neil Wallace of Pennsylvania State University.

In his famous paper "Econometric Policy Evaluation: A Critique," Robert Lucas spurred the rational expectations revolution by presenting a devastating argument against the use of the macroeconometric models used at the time for evaluating policy.[3]

Econometric Policy Evaluation

To understand Lucas's argument, we must first understand how econometric policy evaluation is done. Say, for example, that the Bank of Canada wants to evaluate the potential effects of changes in the overnight interest rate from the existing level of, say, 5%. Using conventional methods, the Bank of Canada economists would feed different overnight interest rate options—say, 4% and 6%—into a computer version of a model. The model would then predict how unemployment and inflation would change under the different scenarios. Then, the policymakers would select the policy that the model predicted would have the most desirable outcomes.

Relying on rational expectations theory, Lucas identified faulty reasoning in this approach if the model did not incorporate rational expectations, as was true for the macroeconometric models used by policymakers at the time: When policies change, public expectations shift as well. For example, if the Bank of Canada raises the overnight interest rate to 6%, this action might change the way the public forms expectations about where interest rates will head in the future. Those changing expectations, as we've seen, can have a real effect on economic behaviour and outcomes. Yet econometric models that do not incorporate rational expectations ignore the effects of changing expectations and thus are unreliable for evaluating policy options.

APPLICATION The Term Structure of Interest Rates

Let's now apply Lucas's argument to a concrete example involving only one equation typically found in econometric models: the term structure equation. The equation relates the long-term interest rate to current and past values of the short-term interest rate. It is one of the most important equations in macroeconometric models because the long-term interest rate, not the short-term rate, is the one believed to have the larger impact on aggregate demand.

In Chapter 6, we learned that the long-term interest rate is related to an average of expected future short-term interest rates. Suppose that in the past, when the short-term rate rose, it quickly fell back down again; that is, any increase was temporary. Because rational expectations theory suggests that any rise in the short-term interest rate is expected to be only temporary, such a rise should have only a minimal effect on the average of expected future short-term rates. Therefore, the rise in the short-term rate should cause the long-term interest rate to rise by only a negligible amount. The term structure relationship as estimated using past data thus indicates that changes in the short-term interest rate will have only a weak effect on the long-term interest rate.

Suppose the Bank of Canada wants to evaluate the effects on the economy of a policy that is likely to raise the short-term interest rate from its existing level of, say, 5% to a higher level of 6% for the foreseeable future. The term structure equation, which has been estimated using past data, will indicate that just a small change in the long-term interest

[3] *Carnegie-Rochester Conference Series on Public Policy* 1 (1976): 19–46.

rate will occur. However, if the public expects that the short-term rate is rising to a permanently higher level, rational expectations theory indicates that people will no longer expect a rise in the short-term rate to be temporary. Instead, when they see the interest rate rise to 6%, they will expect the average of future short-term interest rates to rise substantially, and so the long-term interest rate will rise greatly, not minimally, as suggested by the term structure equation. You can see that using an econometric model to evaluate the likely outcome of the change in Bank of Canada policy can lead to highly misleading information.

The term structure application demonstrates another aspect of the Lucas critique. The effects of a particular policy depend critically on the public's expectations about the policy. If the public expects the rise in the short-term interest rate to be merely temporary, then the response of the long-term interest rate, as we have seen, will be negligible. If, however, the public expects the rise to be more permanent, the response of the long-term rate will be far greater. ***The Lucas critique points out not only that conventional econometric models cannot be used for policy evaluation but also that the public's expectations about a policy will influence the response to that policy.***

The term structure equation discussed here is only one of many equations in the econometric models to which the Lucas critique applies. In fact, Lucas uses examples of consumption and investment equations in his paper. One attractive feature of the term structure example is that it deals with expectations in the financial market, a sector of the economy for which the empirical evidence supporting rational expectations is much stronger. The Lucas critique should also apply, however, to sectors of the economy with regard to which rational expectations theory is more controversial, because the basic principle of the Lucas critique is not that expectations are always rational but rather that the formation of expectations changes when the behaviour of a forecasted variable changes. This less stringent principle is supported by the evidence in sectors of the economy other than financial markets. ◆

25.2 POLICY CONDUCT: RULES OR DISCRETION?

LO 25.2 Compare and contrast the use of policy rules versus discretionary policy.

The Lucas critique exposed the need for new policy models that reflected the insights of rational expectations theory. Here, we explore the implications of the critique on a long-running debate among economists: whether monetary policymakers should have the flexibility to adapt their policies to a changing situation, or whether they should adopt **rules**, binding plans that specify how policy will respond (or not respond) to particular data such as unemployment and inflation data.

Discretion and the Time-Inconsistency Problem

We say that policymakers operate with **discretion** when they make no commitment to future actions but instead make what they believe in that moment to be the right policy decision for the situation. Complexities are introduced when policymakers are empowered to shape policy "on the fly." The *time-inconsistency problem* discussed in Chapter 17 reveals the potential limitations of discretionary policy. Recall that the time-inconsistency problem is the tendency of policymakers to deviate from good long-run plans when making short-run decisions. Policymakers are always tempted to pursue policies that are more expansionary than firms or people expect because such a policy

will boost economic output (and lower unemployment) in the short run. The best policy, however, is *not* to pursue expansionary policy, because decisions about wages and prices reflect workers' and firms' expectations about policy (an implication of the rational expectations revolution). For example, when workers and firms see a central bank pursuing discretionary expansionary policy, they will recognize that this policy is likely to lead to higher inflation in the future. They will therefore raise their expectations of inflation, driving wages and prices up. The rise in wages and prices will lead to higher inflation but may not result in higher output on average.

Policymakers will promote better inflation performance in the long run if they do not try to surprise people with an unexpectedly expansionary policy but instead keep inflation under control. One way of doing this is to abandon discretion and adopt rules to govern policymaking.

Types of Rules

In contrast to discretion, rules are essentially automatic. One famous type of rule, advocated by Milton Friedman and his followers, who are known as **monetarists**, is the **constant-money-growth-rate rule**, whereby policymakers keep the money supply growing at a constant rate regardless of the state of the economy. Other monetarists, such as Bennett McCallum and Alan Meltzer, have proposed variants of this rule that allow the rate of money supply growth to be adjusted for shifts in velocity, nominal income divided by the quantity of money, which has often been found to be unstable in the short run. Rules of this type are deemed nonactivist because they do not "react" to economic conditions. Monetarists advocate rules of this type because they believe that money is the sole source of fluctuations in aggregate demand and that long and variable lags in the effects of monetary policy will lead to greater volatility in economic activity and inflation if policy actively responds to unemployment (as discussed in Chapter 24).

Activist rules, in contrast, specify that monetary policy should react to the level of output as well as to inflation. The most famous rule of this type is the *Taylor rule*, discussed in Chapter 17. It specifies that the Fed should set its federal funds rate target by using a formula that considers both the output gap ($Y - Y^P$) and the inflation gap ($\pi - \pi^T$).

The Case for Rules

As suggested by our discussion of the time-inconsistency problem, discretionary monetary policy can lead to poor economic outcomes. If monetary policymakers operate with discretion, they will be tempted to pursue overly expansionary monetary policies that will boost employment in the short run but generate higher inflation (but not higher employment) in the long run. A commitment to a policy rule like the Taylor rule or the constant-money-growth-rate rule solves the time-inconsistency problem because policymakers are forced to follow a set plan that does not allow them to exercise discretion and therefore prevents them from exploiting the short-run trade-off between inflation and employment. With their hands bound by a policy rule, policymakers can achieve desirable long-run outcomes.

Another argument in support of rules posits that policymakers and politicians simply cannot be trusted to make good decisions. Milton Friedman and Anna Schwartz's monumental work, *A Monetary History of the United States*,[4] documents numerous instances in which the Federal Reserve made serious policy errors, with the worst

[4]Milton Friedman and Anna Jacobson Schwartz, *A Monetary History of the United States, 1867–1960* (Princeton, NJ: Princeton University Press, 1963).

FYI The Political Business Cycle and Richard Nixon

You might know that former U.S. president Richard Nixon and his aides took some extraordinary actions to ensure a landslide victory in the 1972 U.S. presidential election, such as breaking into the offices of political rivals at the Watergate Hotel. Less well known are similar actions on the economic front prior to the election. Before the 1972 election, the Nixon administration imposed wage and price controls on the economy, which temporarily lowered the inflation rate. After the election, those same actions contributed to a surge in inflation. Nixon also pursued expansionary fiscal policy by cutting taxes. And it has been rumoured that the chair of the Federal Reserve at the time, Arthur Burns, succumbed to direct pressure from Nixon to maintain low interest rates through election day. The aftermath was ugly. The economy overheated and inflation rose to over 10% by the late 1970s, a rate that was abetted by the negative supply shocks during that period.

The Nixon episode led economists and political scientists to theorize that politicians will take steps to make themselves look good during election years. Specifically, the theory went, they will take steps to stimulate the economy prior to the election, in the hope of stimulating a boom and low unemployment, outcomes that will increase their electoral chances. Unfortunately, the result of these actions will be higher inflation down the road, which then will require contractionary policies to get inflation under control, leading to a recession in the future. The resulting ups and downs of the economy would be the direct result of political machinations and so could be characterized as a political business cycle. Although the Nixon episode provided support for the existence of a political business cycle, the research has not led to a definitive answer regarding whether this phenomenon is a general one.*

*The paper that launched research on the political business cycle was William Nordhaus, "The Political Business Cycle," *Review of Economic Studies* 42 (1975): 169–190.

occurring during the Great Depression, when the Fed just stood by and let the banking system and the economy collapse. Proponents of this argument believe that the politicians who make fiscal policy are not to be trusted because they have strong incentives to pursue policies that will help them win the next election. They are therefore more likely to focus on increasing employment in the short run without worrying that their actions might lead to higher inflation further down the road. Their advocacy for expansionary policies can lead to a so-called **political business cycle**, in which fiscal and monetary policy are expansionary right before elections, with higher inflation following later. (For an example, see the FYI box "The Political Business Cycle and Richard Nixon.")

The Case for Discretion

Although policy rules have important advantages, they also have serious drawbacks. First, rules can be too rigid because they cannot account for every contingency. For example, almost no one could have predicted that problems in one small part of the financial system, subprime mortgage lending, would lead to the worst financial crisis in over 70 years, with such devastating effects on the world economy. The unprecedented steps taken by the Federal Reserve during this crisis to prevent it from causing a depression could not have been written into a policy rule ahead of time. The ability to act flexibly through discretion can thus be a key factor in a successful monetary policy.

The second problem with policy rules is that they do not easily incorporate the tool of judgment. Monetary policy is as much an art as it is a science. Monetary

policymakers must look at a wide range of information to decide on the best course for monetary policy, and some of this information is not easily quantifiable. Judgment is thus an essential element of good monetary policy, and it is very hard to write it into a rule. Only through discretion can monetary policy bring judgment to bear.

Third, no one really knows what the true model of the economy is, and so any policy rule that is based on a particular model will prove to be wrong if the model is flawed. Discretion avoids the straightjacket that would lock in the wrong policy if the model that was used to derive the policy rule proved to be a poor one.

Fourth, even if the model is a good one, any structural changes in the economy will necessitate changes in the coefficients of the model. For example, the Lucas critique points out that the structural shifts caused by changes in policies can lead to changes in the coefficients of macroeconometric models. Another example of structural economic change occurred in the 1980s, when the relationship between monetary aggregates, such as M1+ and M2+, and aggregate spending broke down as a result of financial innovation. Following a rule that assumed a constant growth rate in one of these aggregates would have produced bad outcomes. Indeed, this is what happened in Switzerland in the late 1980s and early 1990s when adherence to a rule that assumed a particular growth rate of monetary aggregates led to a surge in inflation (as discussed in the Global box "The Demise of Monetary Targeting in Switzerland"). Discretion enables policymakers to change policy settings when an economy undergoes structural changes.

Constrained Discretion

The tug-of-war between rules and discretion has strongly influenced academic debates about monetary policy for many decades. But either choice may be too rigid. As we have seen, both rules and discretion are subject to problems, and so the dichotomy between rules and discretion may be too simple to capture the realities that macroeconomic

Global The Demise of Monetary Targeting in Switzerland

In 1975, the Swiss National Bank (Switzerland's central bank) adopted monetary targeting in which it announced a growth rate target for the monetary aggregate M1. In 1980, the Swiss switched their growth rate target to an even narrower monetary aggregate, the monetary base. Although monetary targeting had been quite successful in Switzerland for many years, the Swiss economy ran into serious problems with the introduction of a new interbank payment system, the Swiss Interbank Clearing, and a wide-ranging revision of commercial banks' liquidity requirements in 1988. These structural changes caused a severe drop in banks' desired holdings of deposits at the Swiss National Bank, and these holdings were the major component of the monetary base. A smaller amount of the monetary base was now needed to conduct transactions, altering the relationship between the monetary base and aggregate spending, and so the 2% target growth rate for the monetary base was suddenly far too expansionary. Inflation subsequently rose to over 5%, well above that in other European countries.

High inflation rates horrified the Swiss, who had always prided themselves on maintaining a low-inflation environment even when the rest of Europe did not. These problems with monetary targeting led the Swiss to abandon it in the 1990s and adopt a much more flexible framework for the conduct of monetary policy.*

*For a further discussion of monetary targeting in Switzerland, see Chapter 4 of Ben S. Bernanke, Thomas Laubach, Frederic S. Mishkin, and Adam S. Posen, *Inflation Targeting: Lessons from the International Experience* (Princeton, NJ: Princeton University Press, 1999).

policymakers face. Discretion can be a matter of degree. Discretion can be a relatively undisciplined approach that leads to policies that change with the personal views of policymakers or with the direction of political winds, or it might operate within a more clearly articulated framework in which the general objectives and tactics of the policymakers—although not their specific actions—are committed to in advance. Ben Bernanke, former chair of the Federal Reserve, along with one of the authors of this textbook, came up with a name for this type of framework: **constrained discretion**.[5] Constrained discretion imposes a conceptual structure and inherent discipline on policymakers, but without eliminating all flexibility. It combines some of the advantages ascribed to rules with those ascribed to discretion.

25.3 THE ROLE OF CREDIBILITY AND A NOMINAL ANCHOR

LO 25.3 Summarize and illustrate the benefits of a credible central bank.

An important way to constrain discretion, as discussed in Chapter 17, is to commit to a *nominal anchor*, a nominal variable—such as the inflation rate, the money supply, or an exchange rate—that ties down the price level or inflation to achieve price stability. For example, if a central bank has an explicit target for the inflation rate—say, 2%—and takes steps to achieve this target, then the inflation target becomes a nominal anchor. Alternatively, a government could commit to a fixed exchange rate between its currency and a sound currency like the U.S. dollar and use this as its nominal anchor. If the commitment to a nominal anchor has **credibility**—that is, it is believed by the public—then it has important benefits.

Benefits of a Credible Nominal Anchor

First, a credible nominal anchor has elements of a behaviour rule. Just as rules help to prevent the time-inconsistency problem in parenting by helping adults resist pursuit of the discretionary policy of giving in, a nominal anchor can help overcome the time-inconsistency problem by providing an expected constraint on discretionary policy. For example, if monetary policymakers commit to a nominal anchor of achieving a specific inflation objective—say, a 2% inflation rate—then they know that they will be subject to public scrutiny and criticism if they miss this objective or pursue policies that are clearly inconsistent with this objective, such as an interest rate target that is too low. Because policymakers wish to avoid embarrassment and possibly legal punishment, they will be less tempted to pursue overly expansionary, discretionary policies in the short run, policies that would be inconsistent with their commitment to the nominal anchor.

Second, a credible commitment to a nominal anchor helps to anchor inflation expectations, which leads to smaller fluctuations in inflation. Such a commitment thus contributes to price stability and also helps stabilize aggregate output. The credibility of a commitment to a nominal anchor is therefore a critical element in enabling monetary policy to achieve both of its objectives, price stability and stabilization of economic activity. In other words, a credible nominal anchor helps make monetary policy more efficient.

[5]See Ben S. Bernanke and Frederic S. Mishkin, "Inflation Targeting: A New Framework for Monetary Policy?" *Journal of Economic Perspectives* 11 (Spring 1997): 97–116.

We will use the aggregate demand and supply framework to show analytically why a credible nominal anchor helps produce this desirable outcome. First we will look at the effectiveness of stabilization policy when it is responding to an aggregate demand shock, and then we will look at its effectiveness when it is responding to an aggregate supply shock. We will also look at the benefits of credibility for anti-inflation policy.

Credibility and Aggregate Demand Shocks

We now examine the importance of credibility in the short run when positive and negative demand shocks are present.

Positive Demand Shock Let's first look at what happens in the short run when a positive aggregate demand shock occurs. For example, suppose businesses suddenly get new information that makes them more optimistic about the future, and so they increase their investment spending. As a result of this positive demand shock, the aggregate demand curve shifts to the right from AD_1 to AD_2, moving the economy from point 1 to point 2 in panel (a) of Figure 25-1. The aggregate output index rises to Y_2 and inflation rises above the inflation target of π^T to π_2. As we saw in Chapter 24, the appropriate response, if the monetary authorities wish to stabilize inflation and economic activity, is to tighten monetary policy and shift the short-run aggregate demand curve back down to AD_1 in order to move the economy back to point 1. However, because of the long lags between monetary policy actions and their effects on aggregate demand, it will take some time before the short-run aggregate demand curve shifts back to AD_1.

Now let's look at what happens to the short-run aggregate supply curve. Recall from Chapter 23 that the short-run aggregate supply curve can be expressed as follows:

$$\begin{array}{ccccccc} \pi & = & \pi^e & + & \gamma(Y - Y^P) & + & \rho \\ \text{inflation} & = & \text{expected inflation} & + & \gamma \times \text{output gap} & + & \text{inflation shock} \end{array}$$

If the commitment to the nominal anchor is credible, then the public's expected inflation rate π^e will remain unchanged and the short-run aggregate supply curve will remain at AS_1. Inflation therefore will not go higher than π_2, and over time, as the short-run aggregate demand curve shifts back down to AD_1, inflation will fall back down to the inflation target of π^T.

But what if monetary policy is not credible? The public will worry that the monetary authorities are willing to accept a higher inflation rate than π^T and are not willing to drive the short-run aggregate demand curve back to AD_1 quickly. In this case, the weak credibility of the monetary authorities will cause expected inflation π^e to rise and so the short-run aggregate supply curve will rise, shifting from AS_1 to AS_3 and sending the economy to point 3 in the short run, where inflation has risen further to π_3. Even if the monetary authorities tighten monetary policy and return the aggregate demand curve to AD_1, the damage is done: Inflation has risen more than it would have if the central bank had credibility. Our aggregate demand and supply analysis thus yields the following conclusion: ***Monetary policy credibility has the benefit of stabilizing inflation in the short run when faced with positive demand shocks.***

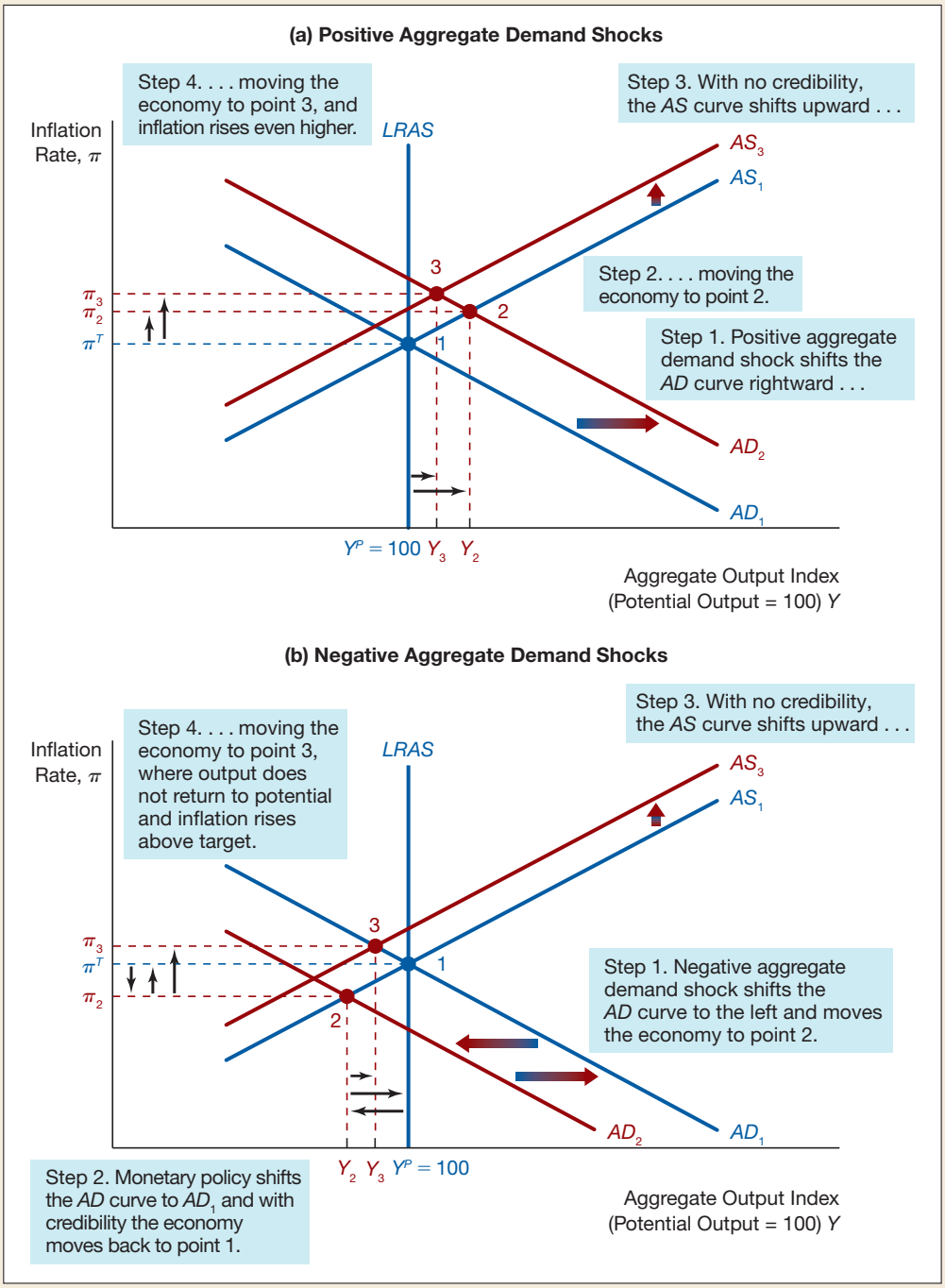
Negative Demand Shock Panel (b) of Figure 25-1 illustrates a negative demand shock. For example, suppose consumer confidence dips and consumer spending declines. The aggregate demand curve shifts left from AD_1 to AD_2, and the economy moves to point 2 in the short run, where the aggregate output index has fallen to Y_2,

Mini-lecture

FIGURE 25-1

Credibility and Aggregate Demand Shocks

In panel (a), the positive aggregate demand shock shifts the aggregate demand curve to the right from AD_1 to AD_2, moving the economy from point 1 to point 2. If monetary policy is not credible, expected inflation will rise, and so the short-run aggregate supply curve will rise and shift to the left to AS_3, sending the economy to point 3, where inflation has risen further to π_3. In panel (b), the negative aggregate demand shock shifts the aggregate demand curve left from AD_1 to AD_2, and the economy moves to point 2. If the central bank is not credible, inflation expectations might increase, and the short-run aggregate supply curve will rise and shift to the left to AS_3. Monetary policy easing then will move the economy to point 3, where the aggregate output index at Y_3 is still below its potential level at $Y^P = 100$.


(a) Positive Aggregate Demand Shocks
Step 4. . . . moving the economy to point 3, and inflation rises even higher.
Step 3. With no credibility, the AS curve shifts upward . . .
Inflation Rate, π
LRAS
AS3
AS1
Step 2. . . . moving the economy to point 2.
Step 1. Positive aggregate demand shock shifts the AD curve rightward . . .
π3
π2
πT
3
2
1
AD2
AD1
YP = 100 Y3 Y2
Aggregate Output Index (Potential Output = 100) Y
(b) Negative Aggregate Demand Shocks
Step 3. With no credibility, the AS curve shifts upward . . .
Step 4. . . . moving the economy to point 3, where output does not return to potential and inflation rises above target.
Inflation Rate, π
LRAS
AS3
AS1
Step 1. Negative aggregate demand shock shifts the AD curve to the left and moves the economy to point 2.
π3
πT
π2
3
1
2
AD2
AD1
Y2 Y3 YP = 100
Step 2. Monetary policy shifts the AD curve to AD1 and with credibility the economy moves back to point 1.
Aggregate Output Index (Potential Output = 100) Y

so output is below potential output, and inflation has fallen to π_2, which is below the target level of π^T. To stabilize output and inflation, the central bank will ease monetary policy to move the aggregate demand curve back to AD_1. If the central bank has high credibility, expected inflation will remain unchanged, the short-run aggregate supply

curve will remain at AS_1, and the economy will return to point 1, where the aggregate output index is back at its potential level Y^P.

However, what if the central bank's credibility is weak? When the public sees an easing of monetary policy, it might become concerned that the central bank is weakening its commitment to the nominal anchor and intends to pursue inflationary policy in the future. In this situation, inflation expectations might increase, and so the short-run aggregate supply curve will rise to AS_3, sending the economy to point 3, where the aggregate output index will rise only from Y_2 to Y_3, still leaving it below the potential level of Y^P. Weak credibility in the face of a negative demand shock will keep output below potential output for a longer period than would occur if credibility was strong. We have the following additional result: ***Monetary policy credibility has the benefit of stabilizing economic activity in the short run when faced with negative demand shocks.***

Credibility and Aggregate Supply Shocks

Now let's look at what happens in Figure 25-2 if a negative aggregate supply shock occurs. If energy prices increase, the short-run aggregate supply curve will shift up and to the left. How much the aggregate supply curve will shift, however, depends on the

Mini-lecture

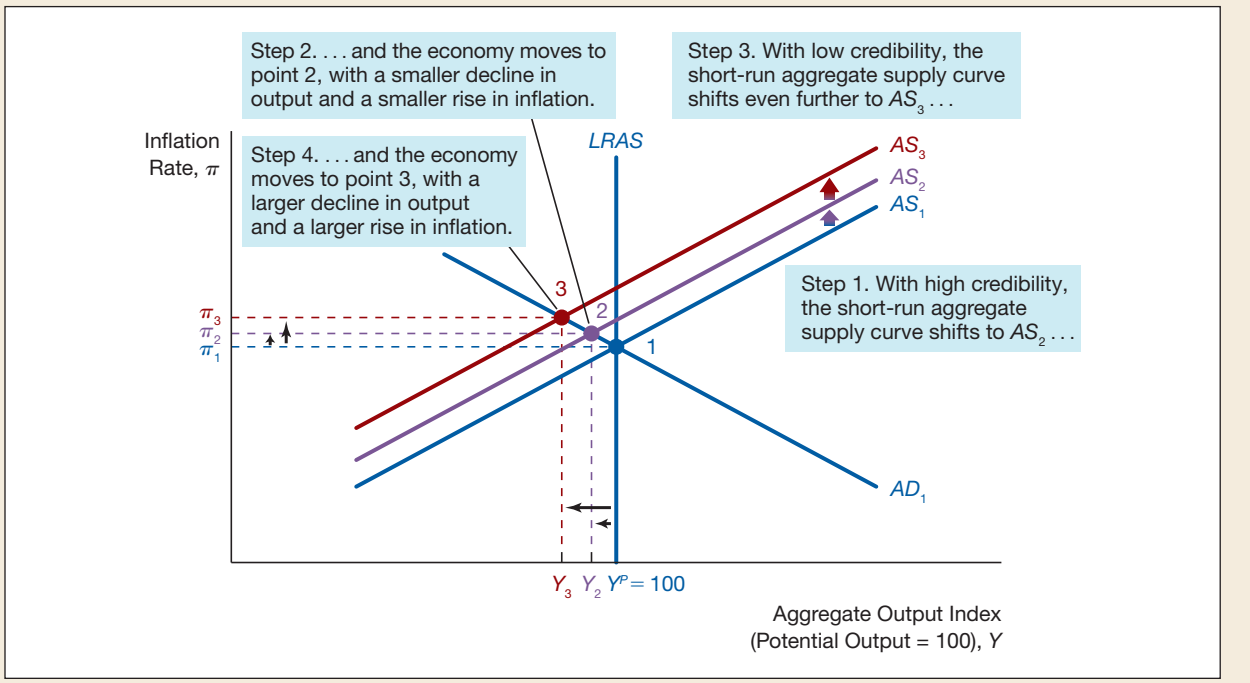

Step 2. . . . and the economy moves to point 2, with a smaller decline in output and a smaller rise in inflation.
Step 3. With low credibility, the short-run aggregate supply curve shifts even further to AS_3 . . .
Inflation Rate, π
Step 4. . . . and the economy moves to point 3, with a larger decline in output and a larger rise in inflation.
LRAS
AS_3
AS_2
AS_1
Step 1. With high credibility, the short-run aggregate supply curve shifts to AS_2 . . .
π_3
π_2
π_1
3
2
1
AD_1
Y_3 Y_2 $Y^P = 100$
Aggregate Output Index (Potential Output = 100), Y

FIGURE 25-2 Credibility and Aggregate Supply Shocks

If the credibility of monetary policy is high, the negative aggregate supply shock will shift the short-run aggregate supply curve only to AS_2 and the economy will move to point 2, where the rise in inflation to π_2 will be minor and the fall in output to Y_2 will also be minor. If credibility is weak, then inflation expectations will rise substantially and the upward shift of the short-run aggregate supply curve will be much larger, moving the curve up and to the left to AS_3. The economy will move to point 3, with worse outcomes for both inflation and output—inflation will be higher at π_3 and the aggregate output index will be lower at Y_3.

amount of credibility the monetary authorities have. If the credibility of the nominal anchor is strong, inflation expectations will not rise, and so the upward and leftward shift of the short-run aggregate supply curve to AS_2 will be small. When the economy moves to point 2 in the short run, the rise in inflation to π_2 will then be minor, and the fall in the aggregate output index to Y_2 will also be minor. If, however, the central bank's commitment to the nominal anchor is perceived as weak, then inflation expectations will rise substantially and the upward and leftward shift of the short-run aggregate supply curve will be much larger, moving the curve up to AS_3. Now the economy will move to point 3 in the short run, with worse outcomes for both inflation and output—inflation will be higher at π_3 and the aggregate output index will be lower at Y_3. We reach the following conclusion: ***Monetary policy credibility has the benefit of producing better outcomes for both inflation and output in the short run when faced with negative supply shocks.***

The theoretical benefits of credibility when the economy is hit by negative supply shocks are confirmed by the data, as illustrated in the following application.

APPLICATION A Tale of Three Oil Price Shocks

In 1973, 1979, and 2007, the world economy was hit by three major negative supply shocks when the price of oil rose sharply; and yet in the first two episodes the inflation rate in Canada rose sharply, while in the most recent episode it rose much less, as we can see in panel (a) of Figure 25-3. In the case of the first two episodes, monetary policy credibility was extremely weak because the Bank of Canada had been unable to keep inflation under control. In contrast, when the third oil price shock hit in 2007–2008, inflation had been low and stable for quite a period of time, and so the Bank of Canada had more credibility regarding its intent to keep inflation under control. One reason that might explain why the third oil price shock appears to have had a smaller effect on inflation is that monetary policy was more credible at that point in time. Our aggregate demand and supply analysis provides the reasoning behind this view.

In the first two episodes, during which both the Bank of Canada's commitment to a nominal anchor and the Bank's credibility were weak, the oil price shocks would have produced a surge in inflation expectations and a large upward and leftward shift in the short-run aggregate supply curve to AS_3 in Figure 25-2. Thus, our aggregate demand and supply analysis predicts that there would be a sharp contraction in economic activity and a sharp rise in inflation. This is exactly what we see in panels (a) and (b) of Figure 25-3. The economic contractions were very severe, with unemployment rising to almost 13% in the aftermath of the 1979 oil price shock (during the 1981–1982 recession). In addition, inflation shot up to around double-digit levels.

In the 2007–2008 episode, the outcome was quite different. Through greater policy credibility established over many years, inflation expectations remained grounded when the oil price shock occurred. As a result, the short-run aggregate supply curve shifted up and to the left by much less, to only AS_2 in Figure 25-2. Our aggregate demand and supply analysis predicts that a much smaller increase in inflation would result from the negative supply shock and that the contraction in economic activity also would be less. Indeed, this is exactly what transpired until the global financial crisis entered its virulent phase in the fall of 2008. Inflation rose by much less than in the previous episodes, and the economy held up fairly well until the financial crisis took a disastrous turn in October 2008. Only then did the economy go into a tailspin, but it is clear that this economic contraction was not the result of the negative supply shock. Inflation actually fell quite dramatically, indicating that a massive negative demand shock was the source of the sharp contraction in economic activity. ◆

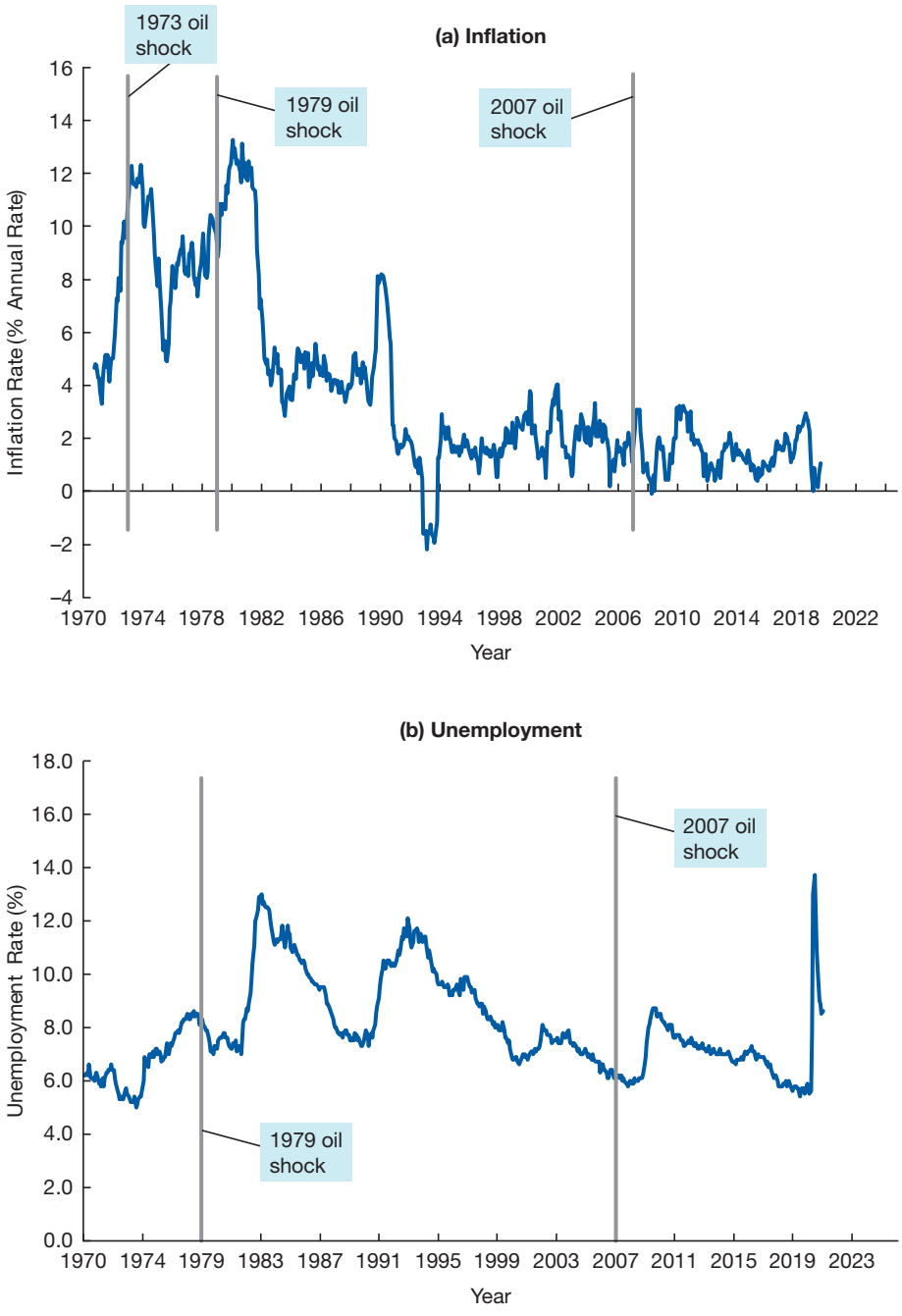


FIGURE 25-3 Inflation and Unemployment, 1970–2020

During the 1973 and 1979 oil shock episodes, inflation was initially high and the Bank of Canada's commitment to a nominal anchor was weak, whereas during the 2007 episode, inflation was initially low and the Bank of Canada's credibility was high. As a result, inflation and unemployment rose more in the first two episodes than in the last episode, with unemployment rising sharply in the third episode only after the 2007–2009 global financial crisis took a disastrous turn in October 2008.

Source: Federal Reserve Bank of St. Louis, FRED database (series CPALCY01CAM661N and LRUNTTTTCAM156S).

Credibility and Anti-Inflation Policy

So far we have examined the benefits of credibility when the economy is buffeted by demand and supply shocks. By the end of the 1970s, the high inflation rate (then over 10%) helped shift the primary concern of policymakers to the reduction of inflation. What role does credibility play in the effectiveness of anti-inflation policies? The aggregate demand and supply diagram given in Figure 25-4 will help us answer this question.

Suppose the economy has settled into a sustained 10% inflation rate at point 1, with the aggregate demand curve at AD_1 and the short-run aggregate supply curve at AS_1. Now suppose a new Bank of Canada governor is appointed who decides that inflation must be stopped. The new governor convinces the Governing Council to autonomously tighten monetary policy so that the aggregate demand curve is shifted to the left to AD_4, which in the long run will move the economy to point 4 and slow inflation down to the 2% inflation objective.

If the central bank has very little credibility, then the public will not be convinced that the central bank will stay the course to reduce inflation, and therefore they will not revise their inflation expectations down from the 10% level. As a result, the short-run aggregate supply curve will remain at AS_1 and the economy will move to point 2. Although inflation will fall to π_2, the aggregate output index will decline to Y_2.

Now consider what will happen if, instead, the central bank has high credibility, and so the public believes that the central bank will do whatever it takes to lower inflation. In this case, expectations of inflation will fall, the short-run aggregate supply curve will fall to AS_3, and the economy will move to point 3. Inflation then will fall even

Mini-lecture

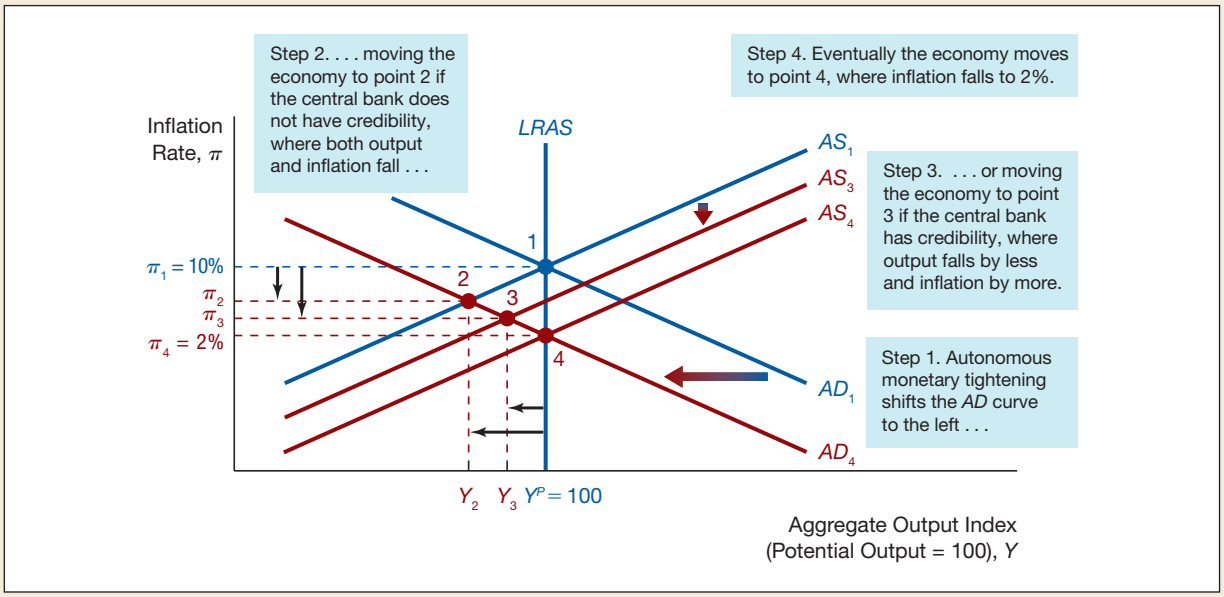

Step 2. . . . moving the economy to point 2 if the central bank does not have credibility, where both output and inflation fall . . .
Step 4. Eventually the economy moves to point 4, where inflation falls to 2%.
Inflation Rate, π
LRAS
AS1
AS3
AS4
Step 3. . . . or moving the economy to point 3 if the central bank has credibility, where output falls by less and inflation by more.
π1 = 10%
π2
π3
π4 = 2%
1
2
3
4
Step 1. Autonomous monetary tightening shifts the AD curve to the left . . .
AD1
AD4
Y2
Y3
YP = 100
Aggregate Output Index (Potential Output = 100), Y

FIGURE 25-4 Credibility and Anti-Inflation Policy

Anti-inflation policy shifts the aggregate demand curve to the left from AD_1 to AD_4. If the central bank lacks credibility, the short-run aggregate supply curve remains at AS_1 and the economy moves to point 2, where the aggregate output index falls to Y_2 and inflation falls to π_2. If the central bank has credibility, then expected inflation declines and the short-run aggregate supply curve shifts down to AS_3. The economy moves to point 3, where the aggregate output index has fallen by less, to Y_3, and inflation has fallen by more, to π_3. Eventually the economy moves to point 4, but with credibility, the anti-inflation policy is more effective, lowering inflation faster and at less cost in terms of output.

further to π_3, and the aggregate output index will decline by less, to Y_3. Indeed, in the extreme case in which the central bank is perfectly credible, there is even a possibility that inflation expectations will immediately decline to the 2% level and that the short-run aggregate supply curve will shift all the way down to AS_4. In this case, the economy will immediately move to point 4, where inflation has declined to the 2% inflation objective, and yet the aggregate output index will remain at Y^P, so the economy stays at full employment. In this case, the anti-inflation policy will be costless because it will not require any loss of output. Although this last scenario is unlikely, we have shown the following result to be true: ***The greater the credibility of the central bank as an inflation fighter, the more rapid the decline in inflation and the lower the loss of output will be in the attempt to achieve the inflation objective.***

Evidence that credibility plays an important role in successful anti-inflation policies is provided by the dramatic end of the Bolivian hyperinflation in 1985 (see the Global box "Ending the Bolivian Hyperinflation: A Successful Anti-Inflation Program"). But establishing credibility is easier said than done. You might think that an announcement by policymakers at the Bank of Canada that they plan to pursue an anti-inflation policy might do the trick. The public will expect that this policy will be adhered to and will act accordingly. However, this conclusion assumes that the public will believe the policymakers' announcement. Unfortunately, that is not how the real world works.

The history of Bank of Canada policymaking suggests that the Bank of Canada has not always done what it set out to do. In the early 1970s, the Bank of Canada accommodated inflationary shocks by raising the rate of growth of monetary aggregates, thereby forming expectations of rising inflation. When in 1975 the Bank of Canada adopted a gradual anti-inflation policy, the public had no reason to believe in such a policy. Such episodes reduced the credibility of the Bank of Canada in the eyes of the public, and, as Figure 25-4 predicts, the resulting lack of credibility had serious consequences. The reduction of inflation that occurred from 1981 to 1984 was bought at a very high cost; the 1981–1982 recession that helped bring the inflation rate down was one of the most severe recessions in the post–World War II period.

Global Ending the Bolivian Hyperinflation: A Successful Anti-Inflation Program

The most remarkable anti-inflation program of recent times was implemented in Bolivia. In the first half of 1985, Bolivia's inflation rate was running at 20 000% and rising. Indeed, the inflation rate was so high that the price of a movie ticket often rose while people waited in line to buy it. In August 1985, Bolivia's new president announced his anti-inflation program, the New Economic Policy. To rein in money growth and establish credibility, the new government took drastic actions to slash the budget deficit by shutting down many state-owned enterprises, eliminating subsidies, freezing public sector salaries, and collecting a new wealth tax. The finance ministry was put on new footing; the budget was balanced on a day-by-day basis. Without exception, the finance minister would not authorize spending in excess of the amount of tax revenue that had been collected the day before. The Bolivian inflation was stopped in its tracks within one month, and the output loss was minor (less than 5% of GDP).

Certain hyperinflations before World War II were also ended with small losses of output through the implementation of policies similar to Bolivia's,* and a 1985 anti-inflation program in Israel, which also involved substantial reductions in budget deficits, sharply reduced inflation without any clear loss of output. Without a doubt, credible anti-inflation policies can be highly successful in eliminating inflation.

*For an excellent discussion of the end of four hyperinflations in the 1920s, see Thomas Sargent, "The Ends of Four Big Inflations," in *Inflation: Causes and Consequences,* ed. Robert E. Hall (Chicago: University of Chicago Press, 1982), 41–98.

FYI Credibility and the Bank of Canada's Victory over Inflation

In the period from 1965 through the 1970s, policymakers in Canada had little credibility as inflation fighters—a well-deserved reputation, as they pursued discretionary accommodating policy to achieve high employment. The outcome was not a happy one. Inflation soared to double-digit levels while the unemployment rate remained high. To wring inflation out of the system, the Bank of Canada under governor Gerald Bouey put the economy through two back-to-back recessions in 1980 and 1981–1982. Only after the 1981–1982 recession—the most severe in the postwar period, with unemployment above the 10% level—did Bouey establish credibility for the Bank of Canada's anti-inflation policy. By the end of 1983, inflation was running at a rate of less than 5%.

From November 1982 to January 1988, the primary objective of the Bank of Canada was price stability in the longer term and inflation containment in the shorter term. This policy, however, was carried out without intermediate targets or a specified path to the longer-term goal. In January 1988, John Crow, the governor of the Bank of Canada, announced that the Bank would subsequently pursue an objective of price stability. One indication of the Bank of Canada's credibility came in 1991 when the Bank's governor and the minister of finance jointly announced a series of declining inflation targets, with a band of plus and minus one percentage point around them. The targets were 3% by the end of 1992, falling to 2% by the end of 1995, to remain within a range of 1% to 3% thereafter. The 1% to 3% target range for inflation has recently been renewed, to apply until the end of 2026.

This has convinced the public and the markets that if inflation reared its head, the Bank of Canada would pursue a nonaccommodating policy of quashing it. Workers and firms have not raised wages and prices, which would have led to both inflation and unemployment. The success of the Bank of Canada's anti-inflation policy has been continuing to date—the inflation rate has been running at a rate of less than 3%. The Bank of Canada's triumph over inflation has been achieved, because it obtained credibility the hard way: It earned it.

25.4 APPROACHES TO ESTABLISHING CENTRAL BANK CREDIBILITY

LO 25.4 Identify the ways in which central banks can establish and maintain credibility.

Our analysis has demonstrated that a credible nominal anchor that anchors inflation expectations is a key element in the success of monetary policy. But how do the monetary authorities achieve this credibility? We discussed several approaches in earlier chapters. One approach to credibility, discussed in Chapter 17, is through continued success at keeping inflation under control through concerted policy actions. An approach that has been growing in popularity in recent years, also discussed in Chapter 17, is *inflation targeting*, which involves the public announcement of a medium-term numerical target for inflation and a commitment to achieving it. Some countries have successfully kept inflation under control by pegging their currency's exchange rate to the currency of an anchor country that already has a strong nominal anchor. We discussed this strategy, sometimes called *exchange-rate targeting*, in Chapter 19. Another way to increase central bank credibility is to give the central bank more independence from the political process, as discussed in Chapter 14. Two additional approaches, not discussed in earlier chapters but discussed below, are nominal GDP targeting and the appointment of "conservative" central bankers.

Nominal GDP Targeting

A variant of inflation targeting that recently has received increased attention is **nominal GDP targeting**, in which the central bank announces an objective of hitting a particular level of nominal GDP growth (recall that nominal GDP is real GDP times the price level). For example, if the inflation objective for the central bank were 2% and potential GDP were expected to grow at an annual rate of 3%, then nominal GDP targeting would mean a commitment by the central bank to keeping nominal GDP growing at 5% per year. Nominal GDP targeting has elements of an inflation targeting regime because the targeted growth rate of nominal GDP is tied to the chosen numerical inflation objective. In addition, nominal GDP targeting implies that the central bank will respond to slowdowns in the real economy even if inflation is not falling. To see this, note that with inflation unchanged, a slowdown in real GDP will result in a slowdown in nominal GDP, and so the monetary authorities will be prompted to pursue more expansionary monetary policy.

A potential advantage of nominal GDP targeting is that it focuses not only on controlling inflation but also explicitly on stabilizing real GDP. Another potential advantage is that with nominal GDP targeting, real GDP growth that is below potential or inflation that is below the inflation objective will encourage more expansionary monetary policy, because actual nominal GDP will fall further below its target. Expectations of this more expansionary policy will help stimulate aggregate demand, which will be particularly useful if the zero-lower-bound problem occurs and the monetary authorities are unable to lower the policy interest rate.

Critics cite two major weaknesses of nominal GDP targeting. First, nominal GDP targeting requires accurate estimates of potential GDP growth, which are not easy to achieve. Second, nominal GDP targeting is more complicated to explain to the public than inflation targeting, and thus the public might be confused about the objectives of the central bank. At this time, no central bank has adopted nominal GDP targeting, but this may change in the future.

Appoint "Conservative" Central Bankers

Kenneth Rogoff of Harvard University suggested that another way of establishing central bank credibility is for the government to appoint central bankers who have a strong aversion to inflation.[6] He characterized these central bankers as "conservative," although a better description would be "tough" or "hawkish on inflation."

When the public sees the appointment of a "conservative" central banker, it will expect that they will be less tempted to pursue expansionary monetary policy to exploit the short-run trade-off between inflation and unemployment and will do whatever it takes to keep inflation under control. As a result, inflation expectations and realized inflation are likely to be more stable, with the benefits outlined previously.

The problem with this approach to solving the credibility problem is that it is not clear that it will continue to work over time. If a central banker has more "conservative" preferences than the public, won't the public demand the appointment of central bankers who are more in tune with their preferences? After all, in a democratic society, government officials are supposed to represent the will of the people.

[6]Kenneth Rogoff, "The Optimal Degree of Commitment to an Intermediary Monetary Target," *Quarterly Journal of Economics* (November 1985): 169–189.

SUMMARY

1. The simple principle (derived from rational expectations theory) that the manner in which expectations are formed changes when the behaviour of forecasted variables changes led to the famous Lucas critique of econometric policy evaluation. Lucas argued that when policy changes, the method of expectations formation changes; hence the relationships in an econometric model will change. An econometric model that has been created on the basis of past data will no longer be a valid model for evaluating the effects of a policy change and may prove to be highly misleading. The Lucas critique also points out that the effects of a particular policy depend critically on the public's expectations of the policy.
2. Advocates of rules for the conduct of monetary policy believe that rules solve the time-inconsistency problem because policymakers are forced to follow a set plan, which eliminates the temptation to deviate from the plan and ensures the achievement of desirable long-run outcomes. Advocates of discretion believe that rules are much too rigid because they cannot predict every contingency and do not allow for the use of judgment. Constrained discretion imposes a conceptual structure and inherent discipline on policymakers, but it does so without eliminating all flexibility so that it combines some of the advantages ascribed to rules with those ascribed to discretion.
3. An important way for the monetary authorities to constrain discretion is by committing to a credible nominal anchor—a nominal variable, such as the inflation rate, the money supply, or an exchange rate—that ties down the price level or inflation to achieve price stability. A credible nominal anchor helps solve the time-inconsistency problem and helps anchor inflation expectations. Credibility has the benefit of stabilizing both output and inflation fluctuations and also enables anti-inflation policies to lower inflation at a lower cost in terms of lost output.
4. Approaches to establishing credibility include implementing actual policies to keep inflation low, inflation targeting, exchange-rate targeting, and the promotion of central bank independence. Two other approaches to establishing central bank credibility are nominal GDP targeting, in which the central bank announces an objective of hitting a particular level of nominal GDP growth, and the appointment of a "conservative" central banker who is hawkish on controlling inflation.

MyLab Economics Make the grade with MyLab Economics: Select end-of-chapter questions can be found on MyLab Economics. You can practise them as often as you want, and most feature step-by-step guided instructions to help you find the right answer.

QUESTIONS

1. What does the Lucas critique state about the limitations of our current understanding of the way in which the economy works?
2. "The Lucas critique casts doubt on the ability of discretionary stabilization policy to be beneficial." Is this statement true, false, or uncertain? Explain your answer.
3. Suppose an econometric model based on past data predicts a small decrease in domestic investment when the Bank of Canada increases the overnight interest rate. Assume the Bank of Canada is considering an increase in the overnight interest rate target to fight inflation and promote a low-inflation environment that will encourage investment and economic growth.
   a. Discuss the implications of the econometric model's predictions if individuals interpret the increase in the overnight interest rate target as a sign that the Bank will keep inflation at low levels in the long run.
   b. What would be Lucas's critique of this model?
4. If the public expects the Bank of Canada to pursue a policy that is likely to raise short-term interest rates permanently to 5%, but the Bank of Canada does not go through with this policy change, what will happen to long-term interest rates? Explain your answer.

5. In what sense can greater central bank independence make the time-inconsistency problem worse?
6. What are the arguments for and against policy rules?
7. If, in a surprise victory, a new administration that the public believes will pursue inflationary policy is elected to office, predict what might happen to the level of output and inflation even before the new administration comes into power.
8. Many economists are worried that a high level of budget deficits may lead to inflationary monetary policies in the future. Could these budget deficits have an effect on the current rate of inflation?
9. In some countries, the government chooses the head of the central bank. The same government can fire the head of the central bank and replace them with another director at any time. Explain the implications of such a situation for the conduct of monetary policy. Do you think the central bank will follow a monetary policy rule, or will it engage in discretionary policy?
10. How would an unexpected change in the equilibrium real overnight interest rate be an argument against using a Taylor rule for monetary policy implementation?
11. How is *constrained discretion* different from *discretion* in monetary policy? How are the outcomes of these policies likely to differ?
12. In general, how does credibility (or lack thereof) affect the aggregate supply curve?
13. In Japan, the government and central bank have enacted policies recently to raise inflation permanently from persistently low levels; however, inflation continues to remain near zero. How, if at all, might credibility of the central bank explain the low inflation persistence?
14. As part of its response to the global financial crisis, the Federal Reserve in the United States lowered the federal funds rate target to nearly zero by December 2008 and quadrupled the monetary base between 2008 and 2014, a considerable easing of monetary policy. However, survey-based measures of 5- to 10-year inflation expectations remained low throughout most of this period. Comment on the Fed's credibility in fighting inflation.
15. "The more credible the policymakers who pursue an anti-inflation policy, the more successful that policy will be." Is this statement true, false, or uncertain? Explain your answer.
16. Why did the oil price shocks of the 1970s affect the economy differently than the oil price shocks of 2007?
17. Central banks that engage in inflation targeting usually announce the inflation target and time period for which that target will be relevant. In addition, central bank officials are held accountable for their actions (e.g., they could be fired if the target is not reached), and their success or lack thereof is also public information. Explain why transparency is such a fundamental ingredient of inflation targeting.
18. In recent years, central banks have dramatically increased the amount of communication with market participants and the public, and at the same time in many of these countries, average inflation has declined and become less volatile. Is this coincidence, or is there a connection? Explain.
19. What are the purposes of inflation targeting, and how does this monetary policy strategy achieve them?
20. How can the establishment of an exchange-rate target bring credibility to a country with a poor record of inflation stabilization?
21. What traits characterize a "conservative" central banker?

APPLIED PROBLEMS

22. Suppose the central bank is following a constant-money-growth-rate rule and the economy is hit with a severe economic downturn. Use an aggregate supply and demand graph to show the possible effects on the economy. How does this situation reflect on the credibility of the central bank if it maintains the money growth rule? How does it reflect on the central bank's credibility if it abandons the money growth rule to respond to the downturn?
23. Suppose country A has a central bank with full credibility and country B has a central bank with no credibility. How does the credibility of each country's central bank affect the speed of adjustment of the aggregate supply curve to policy announcements? How does this result affect output stability? Use an aggregate supply and demand diagram to demonstrate.

24. Suppose two countries have identical aggregate demand curves and potential levels of output and γ is the same in both countries. Assume that in 2022, both countries are hit with the same negative supply shock. Given the table of values below for inflation in each country, what can you say, if anything, about the credibility of each country's central bank? Explain your answer.

	Country A	Country B
2021	3.0%	3.0%
2022	3.8%	5.5%
2023	3.5%	5.0%
2024	3.2%	4.3%
2025	3.0%	3.8%

25. How does a credible nominal anchor help improve the economic outcomes that result from a positive aggregate demand shock? How does a credible nominal anchor help if a negative aggregate supply shock occurs? Use graphs of aggregate supply and demand to demonstrate.

DATA ANALYSIS PROBLEMS

1. Go to the Bank of Canada's web page and find data on the inflation rate (the 12-month growth rate in total CPI).

a. Calculate the average inflation rate over the last 12 and the last 24 months of data available. How does it compare to the 2% inflation goal?

b. What, if anything, does your answer to part (a) imply about Bank of Canada's credibility?

2. **Real-Time Data Analysis** Go to the St. Louis Federal Reserve FRED database and find data on the core PCE price index (PCEPILFE) and the spot price of a barrel of oil (WTISPLC). For both variables, convert the *units* setting to "Percent Change from Year Ago," and download the data from 1960 to the most recent available data.

a. Identify periods in which oil price inflation is 80% or higher.

b. In the periods identified in part (a), how many months was oil price inflation 80% or higher? What was the average core inflation rate during each of those episodes?

c. On the basis of your answers to parts (a) and (b) above, what can you conclude about the credibility of more recent monetary policy compared to its credibility in the earlier periods?

26 Transmission Mechanisms of Monetary Policy

Learning Objectives

26.1 List and summarize the transmission mechanisms through which monetary policy can affect the real economy.

26.2 Summarize and apply the four lessons outlined in this chapter for the conduct of monetary policy.

Preview

Since 1980, the Canadian economy has been on a roller coaster, with output, unemployment, and inflation undergoing drastic fluctuations. At the start of the 1980s, inflation was running at double-digit levels, and the recession of 1980 was followed by one of the shortest economic expansions on record. After a year, the economy plunged into the 1981–1982 recession, with the unemployment rate climbing to over 10%, and only then did the inflation rate begin to come down to below the 5% level. The 1981–1982 recession was followed by a long economic expansion that reduced the inflation rate to around 3% and the unemployment rate to below 8% by early 1990. With Iraq's invasion of Kuwait and a rise in oil prices in the second half of 1990, the Canadian economy again plunged into recession. Subsequent growth in the economy was sluggish at first but eventually sped up, lowering the unemployment rate to below 7%. However, with the onset of the global financial crisis, the economy entered a recession in December 2007, with the unemployment rate rising to over 8.5%. Only in June 2009 did the economy start to recover, making the recession of 2007–2009 the longest recession since World War II. After an expansion that lasted more than 10 years, the coronavirus pandemic led to another recession that started in February 2020, with the unemployment rate rising to 14.7% in May. In light of large fluctuations in aggregate output (reflected in the unemployment rate) and inflation, and the economic instability that accompanies them, policymakers face the following dilemma: What policy or policies, if any, should be implemented to reduce fluctuations in output and inflation in the future?

To answer this question, monetary policymakers must be able to accurately assess the timing and effect of their policies on the economy. To make this assessment, they need to understand the mechanisms through which monetary policy affects the economy. In this chapter, we examine the transmission mechanisms of monetary policy and evaluate the empirical evidence on these mechanisms to better understand the role that monetary policy plays in the economy. We will see that these monetary transmission mechanisms emphasize the link between the financial system (which we studied in the first three parts of this book) and monetary theory, the subject of this part of the book.

26.1 TRANSMISSION MECHANISMS OF MONETARY POLICY

LO 26.1 List and summarize the transmission mechanisms through which monetary policy can affect the real economy.

In this section, we examine the ways in which monetary policy affects aggregate demand and the economy, which are referred to as **transmission mechanisms of monetary policy**. We start with interest-rate channels because they are the key monetary transmission mechanism of the *AD/AS* model developed in Chapters 21, 22, and 23 and applied to monetary policy in Chapters 24 and 25.[1]

Traditional Interest-Rate Channels

The traditional view of the monetary transmission mechanism can be characterized by the following schematic, which shows the effect of an easing of monetary policy accomplished by lowering the real interest rate:

$$r\downarrow \Rightarrow I\uparrow \Rightarrow Y^{ad}\uparrow \tag{1}$$

This schematic shows that an easing of monetary policy leads to a fall in real interest rates $(r\downarrow)$, which in turn lowers the real cost of borrowing, causing a rise in investment spending $(I\uparrow)$, thereby leading to an increase in aggregate demand $(Y^{ad}\uparrow)$.

Although Keynes originally emphasized this channel as operating through businesses' decisions about investment spending, the search for new monetary transmission mechanisms led economists to recognize that consumers' decisions about housing and **consumer durable expenditure** (spending by consumers on durable items such as automobiles and refrigerators) are also investment decisions. Thus the interest-rate channel of monetary transmission outlined in Equation 1 applies equally to consumer spending, and in this case I represents investments in residential housing and consumer durable expenditure.

An important feature of the interest-rate transmission mechanism is its emphasis on the *real* (rather than the nominal) interest rate as the rate that affects consumer and business decisions. In addition, it is often the real *long-term* interest rate (not the real short-term interest rate) that is viewed as having the major impact on spending. How is it that a change in the short-term nominal interest rate induced by a central bank results in a corresponding change in the real interest rate on both short- and long-term bonds? We have already seen that the answer lies in the phenomenon of *sticky prices*—the fact that the aggregate price level adjusts slowly over time, so expansionary monetary policy, which lowers the short-term nominal interest rate, also lowers the short-term real interest rate. The expectations hypothesis of the term structure described in Chapter 6, which states that the long-term interest rate is an average of expected future short-term interest rates, suggests that a lower real short-term interest rate, as long as it is expected to persist, leads to a fall in the real long-term interest rate. These lower real interest rates then lead to increases in business fixed investment, residential housing investment, inventory investment, and consumer durable expenditure, all of which produce the rise in aggregate demand.

The fact that the real interest rate rather than the nominal rate affects spending suggests an important mechanism through which monetary policy can stimulate the

[1]An appendix to this chapter found online examines the empirical evidence on the importance of money to economic fluctuations.

economy, even if nominal interest rates hit a floor of zero (the effective lower bound) during a deflationary episode. With nominal interest rates at a floor of zero, a commitment to future expansionary monetary policy can raise expected inflation (π^e), thereby lowering the real interest rate $(r = i - \pi^e)$ even when the nominal interest rate is fixed at zero and stimulating spending through the interest-rate channel:

$$\pi^e \uparrow \Rightarrow r \downarrow \Rightarrow I \uparrow \Rightarrow Y^{ad} \uparrow \tag{2}$$

This mechanism thus indicates that monetary policy can still be effective even when nominal interest rates have already been driven down to zero by the monetary authorities. Indeed, this mechanism explains why during the global financial and coronavirus crises, the Bank of Canada and other central banks resorted to the nonconventional monetary policy of committing to keep the overnight interest rate at zero (or near zero) for an extended period of time. By so doing, the central banks were trying to keep inflation expectations from falling in order to make sure that real interest rates remained low so as to stimulate the economy. In addition, the commitment to keeping interest rates low for an extended period of time would help lower long-term interest rates, which would also induce greater spending.

Some economists, such as John Taylor of Stanford University, take the position that strong empirical evidence exists for substantial interest-rate effects on consumer and investment spending through the real cost of borrowing, making the interest-rate monetary transmission mechanism a strong one. His position is highly controversial, and many researchers, including Ben Bernanke, former chair of the Federal Reserve, and Mark Gertler of New York University, believe that the empirical evidence does not support strong interest-rate effects that operate through the real cost of borrowing.[2] Indeed, these researchers see the empirical failure of traditional interest-rate monetary transmission mechanisms as having provided the stimulus for the search for other transmission mechanisms of monetary policy.

These other transmission mechanisms fall into two basic categories: those operating through asset prices other than interest rates and those operating through asymmetric information effects on credit markets (the **credit view**). These mechanisms are summarized in the schematic diagram in Figure 26-1.

Other Asset Price Channels

One drawback of the aggregate demand analysis in previous chapters is that it focuses on only one asset price, the interest rate, rather than on many asset prices. In addition to bond prices, two other asset prices receive substantial attention as channels for monetary policy effects: foreign exchange rates and the prices of equities (stocks).

Exchange Rate Effects on Net Exports

With the growing internationalization of economies throughout the world and the advent of flexible exchange rates, more attention has been paid to how monetary policy affects exchange rates, which in turn affect net exports and aggregate demand.

The foreign exchange rate channel also involves interest-rate effects because, as we saw in Chapter 18, when domestic real interest rates fall, domestic dollar assets become less attractive relative to assets denominated in foreign currencies. As a result,

[2]See John Taylor, "The Monetary Transmission Mechanism: An Empirical Framework," *Journal of Economic Perspectives* 9 (Fall 1995): 11–26, and Ben Bernanke and Mark Gertler, "Inside the Black Box: The Credit Channel of Monetary Policy Transmission," *Journal of Economic Perspectives* 9 (Fall 1995): 27–48.

Mini-lecture

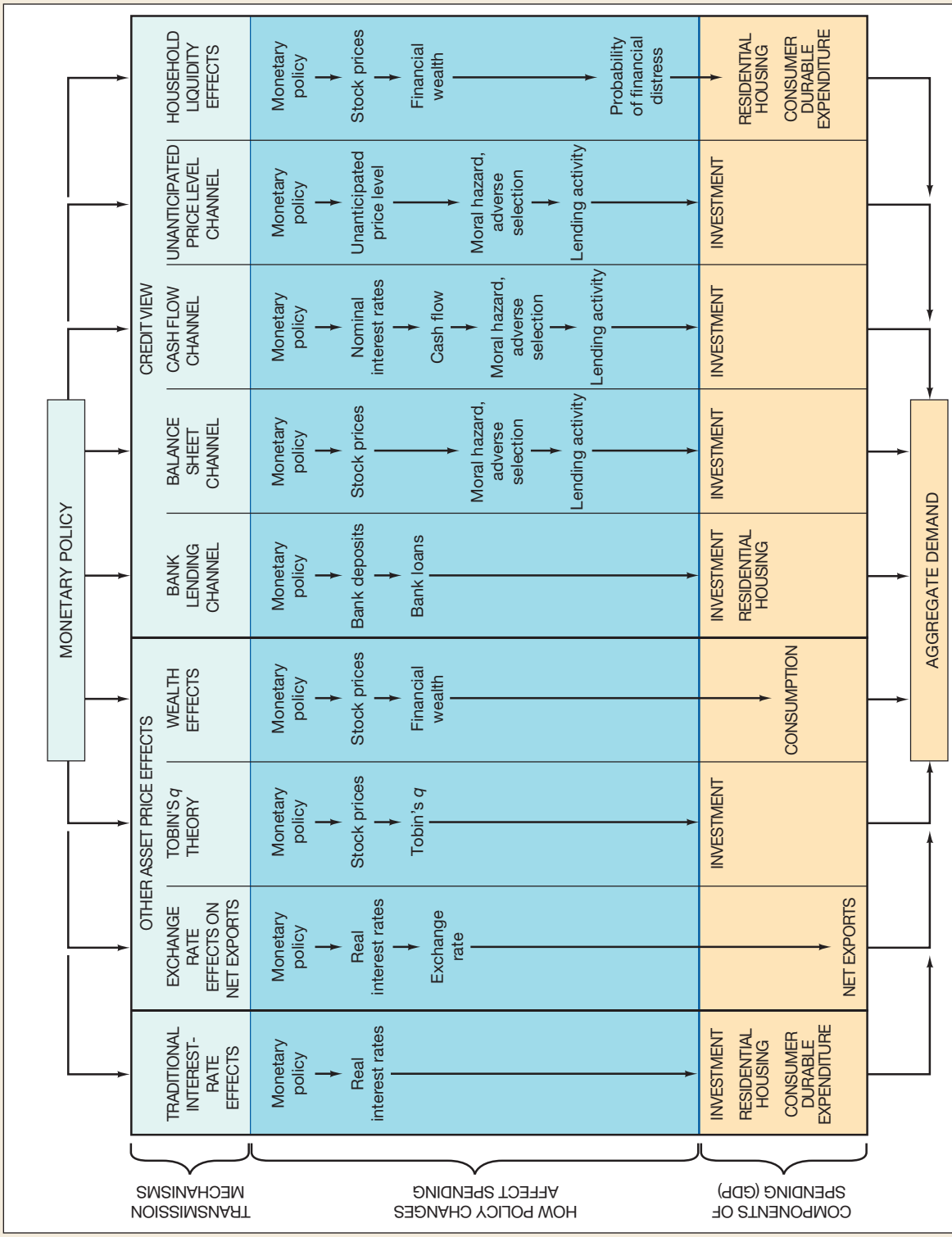


FIGURE 26-1 The Link Between Monetary Policy and Aggregate Demand: Monetary Transmission Mechanisms

This figure shows the different channels through which monetary policy affects aggregate demand.

the value of dollar assets relative to other currency assets falls, and the dollar depreciates (denoted by $E\downarrow$). The lower value of the domestic currency makes domestic goods cheaper than foreign goods, thereby causing a rise in net exports $(NX\uparrow)$ and hence in aggregate demand $(Y^{ad}\uparrow)$. The schematic for the monetary transmission mechanism that operates through the exchange rate is

$$r\downarrow \Rightarrow E\downarrow \Rightarrow NX\uparrow \Rightarrow Y^{ad}\uparrow \tag{3}$$

Tobin's *q* Theory Nobel Prize winner James Tobin developed a theory, referred to as *Tobin's q theory*, that explains how monetary policy can affect the economy through its effects on the valuation of equities (stock). Tobin defines q as the market value of firms divided by the replacement cost of capital. If q is high, the market price of firms is high relative to the replacement cost of capital, and new plant and equipment capital is cheap relative to the market value of firms. Companies then can issue stock and get a high price for it relative to the cost of the facilities and equipment they are buying. Investment spending will rise because firms can buy a lot of new investment goods with only a small issue of stock.

Conversely, when q is low, firms will not purchase *new* investment goods because the market value of firms is low relative to the cost of capital. If companies want to acquire capital when q is low, they can buy another firm cheaply and acquire old capital instead. Investment spending, the purchase of new investment goods, will then be very low. Tobin's q theory gives a good explanation for the extremely low rate of investment spending during the Great Depression. In that period, stock prices collapsed, and by 1933, stocks were worth only one-tenth of their value in late 1929; q fell to unprecedentedly low levels.

The crux of this discussion is that a link exists between Tobin's q and investment spending. But how might monetary policy affect stock prices? Quite simply, lower real interest rates on bonds mean that the expected return on this alternative to stocks falls. This makes stocks more attractive relative to bonds, and so demand for them increases, which raises their price.[3] By combining this result with the fact that higher stock prices (P_s) will lead to a higher q and thus higher investment spending I, we can write the following transmission mechanism of monetary policy:

$$r\downarrow \Rightarrow P_s\uparrow \Rightarrow q\uparrow \Rightarrow I\uparrow \Rightarrow Y^{ad}\uparrow \tag{4}$$

Wealth Effects In their search for new monetary transmission mechanisms, researchers also looked at how consumers' balance sheets might affect their spending decisions. Franco Modigliani was the first to take this tack, using his famous life cycle hypothesis of consumption. **Consumption** is spending by consumers on nondurable goods and services.[4] It differs from *consumer expenditure* in that it does not include spending on consumer durables. The basic premise of Modigliani's theory is that consumers smooth out their consumption over time. Therefore, consumption spending is determined by the lifetime resources of consumers, not just today's income.

[3]An alternative way of looking at this transmission mechanism is to use the model discussed in Chapter 7 in which a decrease in the real interest rate lowers the required return on investments in stocks and so increases stock prices. Then the lower yield on stocks reduces the cost of financing investment spending through the issuance of equity. This way of looking at the link between stock prices and investment spending is formally equivalent to Tobin's q theory.

[4]Consumption also includes another small component, the services that a consumer receives from the ownership of housing and consumer durables.

An important component of consumers' lifetime resources is their financial wealth, a major part of which is common stocks. When stock prices rise, the value of financial wealth increases, thereby increasing the lifetime resources of consumers, which means that consumption should rise. Considering that, as we have seen, monetary easing can lead to a rise in stock prices, we now have another monetary transmission mechanism:

$$r\downarrow \Rightarrow P_s\uparrow \Rightarrow \text{wealth}\uparrow \Rightarrow \text{consumption}\uparrow \Rightarrow Y^{ad}\uparrow \tag{5}$$

Modigliani's research found this relationship to be an extremely powerful mechanism that adds substantially to the potency of monetary policy.

The wealth and Tobin's q channels allow for a general definition of equity, so they can also be applied to the housing market, where housing is equity. An increase in home prices, which raises their prices relative to replacement cost, leads to a rise in Tobin's q for housing, thereby stimulating its production. Similarly, housing prices are extremely important components of wealth, so rises in these prices increase wealth, thereby increasing consumption. Monetary expansion, which raises housing prices through the Tobin's q and wealth mechanisms described here, thus leads to a rise in aggregate demand.

Credit View

Dissatisfaction with the conventional story that interest-rate effects explain the impact of monetary policy on spending on durable assets has led to a new explanation that is based on the concept of asymmetric information, a problem that leads to financial frictions in financial markets (see Chapter 8). This explanation, referred to as the *credit view*, proposes that two types of monetary transmission channels arise as a result of financial frictions in credit markets: those that operate through effects on bank lending and those that operate through effects on firms' and households' balance sheets.

Bank Lending Channel

The concept of the bank lending channel is based on the analysis in Chapter 8, which demonstrated that banks play a special role in the financial system because they are especially well suited to solving asymmetric information problems in credit markets. Because of banks' special role, certain borrowers will not have access to the credit markets unless they borrow from banks. As long as there is no perfect substitutability of retail bank deposits with other sources of funds, the bank lending channel of monetary transmission operates as follows: Expansionary monetary policy, which increases bank reserves and bank deposits, raises the quantity of bank loans available. Because many borrowers are dependent on bank loans to finance their activities, this increase in loans causes investment (and possibly consumer) spending to rise. Schematically, the monetary policy effect is written as follows:

$$\text{bank reserves}\uparrow \Rightarrow \text{bank deposits}\uparrow \Rightarrow \text{bank loans}\uparrow \Rightarrow I\uparrow \Rightarrow Y^{ad}\uparrow \tag{6}$$

An important implication of the credit view is that monetary policy will have a greater effect on expenditure by smaller firms, which are more dependent on bank loans, than it will on large firms, which can get funds directly through the stock and bond markets (and not only through banks).

Balance Sheet Channel

Even though the bank lending channel may be declining in importance, it is by no means clear that this is the case for the other credit channel, the balance sheet channel. Like the bank lending channel, the balance sheet

channel arises from the presence of financial frictions in credit markets. In Chapter 8, we saw that the lower the net worth of business firms, the more severe the adverse selection and moral hazard problems in lending to these firms become. Lower net worth means that lenders in effect have less collateral for their loans, so their potential losses from adverse selection are higher. A decline in firms' net worth, which raises the adverse selection problem, thus leads to decreased lending to finance investment spending. The lower net worth of businesses also increases the moral hazard problem because it means that owners have a lower equity stake in their firms, giving them more incentive to engage in risky investment projects. When borrowers take on more risky investment projects, it is more likely that lenders will not be paid back, and so a decrease in businesses' net worth leads to a reduction in lending and hence in investment spending.

Monetary policy can affect firms' balance sheets in several ways. Easing of monetary policy, which causes a rise in stock prices $(P_s\uparrow)$ along the lines described earlier, raises the net worth of firms and so leads to higher investment spending $(I\uparrow)$ and higher aggregate demand $(Y^{ad}\uparrow)$ because of the decrease in adverse selection and moral hazard problems. This leads to the following schematic for this particular balance sheet channel of monetary transmission:

$$\begin{aligned} r\downarrow \Rightarrow P_s\uparrow \Rightarrow \text{firms' net worth}\uparrow \Rightarrow \text{adverse selection}\downarrow, \\ \text{moral hazard}\downarrow \Rightarrow \text{lending}\uparrow \Rightarrow I\uparrow \Rightarrow Y^{ad}\uparrow \end{aligned} \tag{7}$$

Cash Flow Channel Another balance sheet channel operates by affecting *cash flow*, the difference between firms' cash receipts and cash expenditures. An easing of monetary policy, which lowers nominal interest rates, also causes an improvement in firms' balance sheets because it raises cash flow. The increase in cash flow increases the liquidity of the firm (or household) and thus makes it easier for lenders to know whether the firm (or household) will be able to pay its bills. The result is that adverse selection and moral hazard problems become less severe, leading to an increase in lending and economic activity. The following schematic describes this alternative balance sheet channel:

$$\begin{aligned} i\downarrow \Rightarrow \text{firms' cash flow}\uparrow \Rightarrow \text{adverse selection}\downarrow, \\ \text{moral hazard}\downarrow \Rightarrow \text{lending}\uparrow \Rightarrow I\uparrow \Rightarrow Y^{ad}\uparrow \end{aligned} \tag{8}$$

An important feature of this transmission mechanism is that *nominal* interest rates affect firms' cash flow. Thus this interest-rate mechanism differs from the traditional interest-rate mechanism discussed earlier in which the real interest rate affects investment. Furthermore, the short-term interest rate plays a special role in this transmission mechanism because interest payments on short-term (rather than long-term) debt typically have the greatest impact on the cash flow of households and firms.

A related transmission mechanism involving adverse selection is the credit-rationing phenomenon. Through this mechanism, expansionary monetary policy that lowers interest rates can stimulate aggregate demand. As discussed in Chapter 12, credit rationing occurs when borrowers are denied loans even though they are willing to pay a higher interest rate. The loans are denied because individuals and firms with the riskiest investment projects are exactly the ones who are willing to pay the highest interest rates because, if the high-risk investment succeeds, they will be the primary beneficiaries. Thus higher interest rates increase the adverse selection problem, and lower

interest rates reduce it. When expansionary monetary policy lowers interest rates, risk-prone borrowers make up a smaller fraction of those demanding loans, and so lenders are more willing to lend, raising both investment and aggregate demand, along the lines of parts of the schematic given in Equation 8.

Unanticipated Price Level Channel A third balance sheet channel operates through monetary policy effects on the general price level. Because in industrialized countries debt payments are contractually fixed in nominal terms, an unanticipated rise in the price level lowers the value of firms' liabilities in real terms (decreases the burden of the debt) but should not lower the real value of the firms' assets. An easing of monetary policy, which raises inflation and hence leads to an unanticipated rise in the price level $(P\uparrow)$, therefore raises real net worth, which lowers adverse selection and moral hazard problems, thereby leading to a rise in investment spending and aggregate demand, as in the following schematic:

$$\begin{aligned} r\downarrow \Rightarrow \pi\uparrow \Rightarrow \text{unanticipated } P\uparrow \Rightarrow \text{ firms' real net worth}\uparrow \\ \Rightarrow \text{adverse selection}\downarrow, \text{ moral hazard}\downarrow \Rightarrow \text{lending}\uparrow \Rightarrow I\uparrow \Rightarrow Y^{ad}\uparrow \end{aligned} \tag{9}$$

The view that unanticipated movements in the price level affect aggregate demand has a long tradition in economics: It is the key feature in the debt-deflation view of the Great Depression, outlined in Chapter 11.

Household Liquidity Effects Although most literature on the credit channel focuses on spending by businesses, the credit view should apply equally well to consumer spending, particularly spending on consumer durables and housing. Declines in bank lending induced by a monetary contraction should cause corresponding declines in durable and housing purchases by consumers who do not have access to other sources of credit. Similarly, increases in interest rates should cause deteriorations in household balance sheets, because consumers' cash flow is adversely affected.

The balance sheet channel also operates through liquidity effects on consumer durable and housing expenditures. These effects were found to be important factors during the Great Depression (see the FYI box "Consumers' Balance Sheets and the Great Depression"). In the liquidity effects view, balance sheet effects work through their impact on consumers' desire to spend rather than on lenders' desire to lend. Because of asymmetric information regarding their quality, consumer durables and housing are very illiquid assets. If, as a result of a severe income shock, consumers needed to sell their consumer durables or housing immediately to raise money, they would expect to suffer a big financial loss because they would not be able to get the full value of these assets in a distress sale. (This is just a manifestation of the lemons problem described in Chapter 8.) In contrast, if consumers held financial assets (such as money in the bank, stocks, or bonds), they could sell them quickly and easily for their full market value and raise the cash. Hence, if consumers expect that they are likely to find themselves in financial distress, they will prefer to hold fewer illiquid consumer durable and housing assets and a greater amount of liquid financial assets.

A consumer's balance sheet should be an important influence on their estimate of the likelihood of future suffering from financial distress. Specifically, when consumers have a large amount of financial assets relative to their debts, their estimate of the probability of financial distress is low, and they are more willing to purchase consumer durables or housing. When stock prices rise, the value of financial assets increases as well; consumer durable expenditure will also rise because consumers have a more

secure financial position and therefore a lower estimate of the likelihood of future financial distress. This leads to another transmission mechanism for monetary policy, one that operates through the link between money and stock prices:

$$\begin{aligned} r\downarrow \Rightarrow P_s\uparrow \Rightarrow &\text{ value of households' financial assets}\uparrow \\ \Rightarrow &\text{ likelihood of financial distress}\downarrow \\ \Rightarrow &\text{ consumer durable and housing expenditure}\uparrow \Rightarrow Y^{ad}\uparrow \end{aligned} \tag{10}$$

The illiquidity of consumer durable and housing assets provides another reason why a monetary easing, which lowers interest rates and thereby increases cash flow to consumers, leads to a rise in spending on consumer durables and housing. An increase in consumer cash flow decreases the likelihood of financial distress, which increases the desire of consumers to hold durable goods and housing, thus increasing spending on these items and hence increasing aggregate demand. The only difference between this view of cash flow effects and that outlined in Equation 8 is that in this view, it is not the willingness of lenders to lend to consumers that causes expenditure to rise but the willingness of consumers to spend.

Why Are Credit Channels Likely to Be Important?

There are three reasons to believe that credit channels are important monetary transmission mechanisms. First, a large body of evidence on the behaviour of individual firms supports the view that financial frictions of the type crucial to the operation of credit channels do affect firms' employment and spending decisions. Second, evidence shows that small firms (which are more likely to be credit-constrained) are hurt more by tight monetary policy than large firms, which are unlikely to be credit-constrained. Third, and maybe most compelling, the asymmetric information view of financial frictions, which is at the core of credit channel analysis, is a theoretical construct that has proved useful in explaining many other important economic phenomena, such as why many of our financial institutions exist, why our financial system has the structure that it has, and why financial crises are so damaging to the economy (topics discussed in Chapters 8 and 11). The best support for a theory is its demonstrated usefulness in a wide range

FYI Consumers' Balance Sheets and the Great Depression

The years between 1929 and 1933 witnessed the worst deterioration in consumers' balance sheets ever seen in the United States. The stock market crash in 1929, which caused an economic slump that lasted until 1933, reduced the value of consumers' wealth by $1020 billion (in 2009 dollars), and as expected, consumption dropped sharply (by $199 billion). Because of the decline in the price level during that period, the level of real debt that consumers owed also increased sharply (by over 20%). Consequently, the value of financial assets relative to the amount of debt declined sharply, increasing the likelihood of financial distress. Not surprisingly, spending on consumer durables and housing fell precipitously: From 1929 to 1933, consumer durable expenditure declined by over 50%, while expenditure on housing declined by 80%.*

*For further discussion of the effect of consumers' balance sheets on spending during the Great Depression, see Frederic S. Mishkin, "The Household Balance Sheet and the Great Depression," *Journal of Economic History* 38 (1978): 918–937.

of applications. By this standard, the asymmetric information theory, which supports the existence of credit channels as an important monetary transmission mechanism, has much to recommend it.

APPLICATION The Great Recession

With the advent of the financial crisis in the summer of 2007, the Bank of Canada and many other central banks around the world began a very aggressive easing of monetary policy. For example, the Federal Reserve dropped the target federal funds rate from 5¼% to 0% over a 15-month period from September 2007 to December 2008. At first, it appeared that the Fed's actions would keep the growth slowdown mild and prevent a recession. However, the U.S. economy proved to be weaker than the Fed or private forecasters expected, with the most severe recession of the post–World War II period up until that time beginning in December 2007. Why did the U.S. economy become so weak despite this unusually rapid reduction in the Fed's policy instrument?

The financial meltdown led to negative effects on the economy from many of the channels we have outlined above. The rising level of subprime mortgage defaults, which led to a decline in the value of mortgage-backed securities and CDOs, led to large losses on the balance sheets of financial institutions. With weaker balance sheets, these financial institutions began to deleverage and cut back on their lending. With no one else available to collect information and make loans, adverse selection and moral hazard problems, and hence financial frictions, increased in credit markets, leading to a slowdown of the economy. Credit spreads also went through the roof with the increase in uncertainty caused by the failure of so many financial markets. The declines in the stock market and housing prices also weakened the economy because they lowered household wealth. The decrease in household wealth led to a drop in Tobin's q, which led to restrained consumer spending and weaker investment spending.

With all these channels operating, it is no surprise that despite the Fed's aggressive lowering of the federal funds rate, the U.S. economy still took a big hit. ◆

26.2 LESSONS FOR MONETARY POLICY

LO 26.2 Summarize and apply the four lessons outlined in this chapter for the conduct of monetary policy.

What useful lessons regarding the appropriate conduct of monetary policy can we draw from the analysis in this chapter? Four basic lessons can be learned.

1. ***It is dangerous to consistently associate an easing or tightening of monetary policy with a fall or rise in short-term nominal interest rates.*** Because most central banks use short-term nominal interest rates—typically, the interbank rate—as the key operating instrument in their monetary policies, the danger exists that central banks

and the public will focus too much on short-term nominal interest rates as an indicator of the stance of monetary policy. Indeed, it is quite common to see statements that always associate monetary tightenings with a rise in the interbank rate and monetary easings with a decline in the interbank rate. We do not make this mistake in this book because we have been careful to associate monetary easing or tightening with changes in *real* and not *nominal* interest rates.

2. ***Other asset prices besides those on short-term debt instruments contain important information about the stance of monetary policy because they are important elements in various monetary policy transmission mechanisms.*** As we have seen in this chapter, economists have come a long way in understanding that other asset prices besides interest rates have major effects on aggregate demand. As we saw in Figure 26-1, asset prices such as stock prices, foreign exchange rates, and housing prices play an important role in monetary transmission mechanisms. Furthermore, additional channels, such as those that operate through the exchange rate, Tobin's q, and wealth effects, provide additional evidence that other asset prices play an important role in monetary transmission mechanisms. Although economists strongly disagree among themselves about which channels of monetary transmission are the most important—not surprising, given that economists, particularly those in academia, always enjoy a good debate—they do concur that asset prices other than those on short-term debt instruments play an important role in the effects of monetary policy on the economy.

   The view that asset prices other than short-term interest rates matter has important implications for monetary policy. When we try to assess the stance of policy, it is critical that we look at other asset prices in addition to short-term interest rates. For example, if short-term interest rates are low or even zero and yet stock prices are low, housing prices are low, and the value of the domestic currency is high, monetary policy is clearly tight, not easy.

3. ***Monetary policy can be effective in reviving a weak economy even if short-term interest rates are already near zero.*** We have recently entered a world in which inflation is not always the norm. Japan, for example, recently experienced a period of deflation during which the price level was actually falling. In the United States, the federal funds rate hit a floor of zero from the end of 2008 until the end of 2015 and then again in March 2020. One common view is that when a central bank has driven down short-term nominal interest rates to nearly zero, monetary policy can do nothing more to stimulate the economy. The transmission mechanisms of monetary policy described here indicate that this view is false. As indicated in our discussion of the factors that affect the monetary base in Chapter 15, expansionary monetary policy aimed at increasing liquidity in the economy can be conducted through open market purchases, which do not have to be solely in short-term government securities. For example, purchases of private securities, as the Bank of Canada made in 2009 and in 2020, can reduce financial frictions by lowering credit spreads and stimulating investment spending. In addition, a commitment to future expansionary monetary policy helps revive the economy by raising inflation expectations and by reflating other asset prices, which then stimulate aggregate demand through the channels outlined here. The nonconventional monetary policies we discussed in Chapter 16 are policies of this type. Nonconventional monetary policies can be a potent force in reviving economies that are undergoing deflation and that have short-term interest rates near zero. Indeed, as we saw in Chapter 11, aggressive nonconventional monetary policy during the recent financial crisis helped prevent the Great Recession from turning into a Great Depression and also helped the economy avoid a deflationary episode like the one that occurred during the Great Depression era.

Aggressive nonconventional monetary policy has again been implemented to boost the economy in the aftermath of the coronavirus pandemic.

4. ***Avoiding unanticipated fluctuations in the price level is an important objective of monetary policy, thus providing a rationale for price stability as the primary long-run goal of monetary policy.*** As we saw in Chapter 17, central banks in recent years have been placing greater emphasis on price stability as the primary long-run goal of monetary policy. Several rationales for this goal have been proposed, including the undesirable effects of uncertainty about the future price level on business decisions and hence on productivity, distortions associated with the interaction of nominal contracts and the tax system with inflation, and increased social conflict stemming from inflation. Our discussion of monetary transmission mechanisms provides an additional reason why price stability is so important. As we have seen, unanticipated movements in the inflation rate can cause unanticipated fluctuations in output, an undesirable outcome. Particularly important in this regard is the knowledge that, as we saw in Chapter 11, price deflation can be an important factor leading to a prolonged financial crisis, as occurred during the Great Depression. An understanding of the monetary transmission mechanisms thus makes it clear that the goal of price stability is desirable because it reduces uncertainty about the future price level. The price stability goal implies that a negative inflation rate is at least as undesirable as too high an inflation rate. Indeed, because of the threat of financial crises, central banks must work very hard to prevent price deflation.

APPLICATION Applying the Monetary Policy Lessons to Japan's Two Lost Decades

Until 1990, it looked as if Japan might overtake the United States in per capita income. From the early 1990s until 2012, during the years that have become known as the "two lost decades," the Japanese economy stagnated, with deflation and low growth. As a result, Japanese living standards fell further and further behind those of the United States. Many economists take the view that Japanese monetary policy is in part to blame for the poor performance of the Japanese economy during this period. Could Japanese monetary policy have performed better if Japan had applied the four lessons outlined in the previous section?

The first lesson suggests that it is dangerous to think that declines in interest rates always mean that monetary policy has been easing. In the mid-1990s, when short-term interest rates began to decline, falling to nearly zero in the late 1990s and early 2000s, the monetary authorities in Japan took the view that monetary policy was sufficiently expansionary. Now it is widely recognized that this view was incorrect, because the falling and eventually negative inflation rates in Japan meant that real interest rates were actually quite high and that monetary policy was tight, not easy. If the monetary authorities in Japan had followed the advice of the first lesson, they might have pursued a more expansionary monetary policy, which would have helped boost the economy.

The second lesson suggests that monetary policymakers should pay attention to other asset prices, in addition to those on short-term debt instruments, in assessing the stance of monetary policy. At the same time that interest rates were falling in Japan, stock and real estate prices were collapsing, thus providing another indication that Japanese monetary policy was not easy. Knowledge of the second lesson might have led Japanese monetary policymakers to recognize sooner that they needed a more expansionary monetary policy.

The third lesson indicates that monetary policy can still be effective even if short-term interest rates are near zero. Officials at the Bank of Japan frequently claimed that they were helpless to stimulate the economy because short-term interest rates had fallen to nearly zero. By recognizing that monetary policy can be effective even when interest rates are near zero, as suggested by the third lesson, the Japanese monetary authorities could have taken monetary policy actions that would have stimulated aggregate demand by raising other asset prices and inflationary expectations.

The fourth lesson indicates that unanticipated fluctuations in the price level should be avoided. If the Japanese monetary authorities had adhered to this lesson, they might have recognized that allowing deflation to occur could be very damaging to the economy and that such deflation was inconsistent with the goal of price stability.

These four lessons of monetary policy were finally taken to heart by the Bank of Japan. As we discussed in Chapter 24, monetary policy in Japan has undergone a dramatic shift since 2013, to a highly expansionary, nonconventional monetary policy with a higher inflation target. Since then the Japanese economy has improved, with the unemployment rate falling. However, although the inflation rate has risen, it still remains stubbornly low, well below the Bank of Japan's 2% inflation target. ◆

SUMMARY

1. The transmission mechanisms of monetary policy include traditional interest-rate channels that operate through the real cost of borrowing and affect investment; other asset price channels such as exchange rate effects, Tobin's *q* theory, and wealth effects; and the credit view channels—the bank lending channel, the balance sheet channel, the cash flow channel, the unanticipated price level channel, and household liquidity effects.
2. Four lessons for monetary policy can be drawn from this chapter: (1) It is dangerous to consistently associate monetary policy easing or tightening with a fall or rise in short-term nominal interest rates; (2) other asset prices besides those on short-term debt instruments contain important information about the stance of monetary policy because they are important elements in the monetary policy transmission mechanisms; (3) monetary policy can be effective in reviving a weak economy even if short-term interest rates are already near zero; and (4) avoiding unanticipated fluctuations in the price level is an important objective of monetary policy, thus providing a rationale for price stability as the primary long-run goal of monetary policy.

MyLab Economics Make the grade with MyLab Economics: Select end-of-chapter questions can be found on MyLab Economics. You can practise them as often as you want, and most feature step-by-step guided instructions to help you find the right answer.

QUESTIONS

1. From 2008 to 2015, auto loan rates in the United States declined from around 8% to near-historic lows of around 4%. At the same time, auto sales increased dramatically. How, if at all, does this relate to the monetary transmission mechanisms?
2. "Considering that consumption accounts for nearly two-thirds of total GDP, this means that the interest rate, wealth, and household liquidity channels are the most important monetary policy channels in Canada." Is this statement true, false, or uncertain? Explain your answer.

3. How can the interest-rate channel still function when short-term nominal interest rates are at the effective lower bound?
4. Lars Svensson, a former Princeton professor and deputy governor of the Swedish central bank, proclaimed that when an economy is at risk of falling into deflation, central bankers should be "responsibly irresponsible" with monetary expansion policies. What does this mean, and how does it relate to the monetary transmission mechanisms?
5. Describe an advantage and a disadvantage of the fact that monetary policy has so many different channels through which it can operate.
6. "If countries fix their exchange rate, the exchange rate channel of monetary policy does not exist." Is this statement true, false, or uncertain? Explain your answer.
7. During the 2007–2009 recession, the value of common stocks in real terms fell by more than 50%. How might this decline in the stock market have affected aggregate demand and thus contributed to the severity of the recession? Be specific about the mechanisms through which the stock market decline affected the economy.
8. "The costs of financing investment are related only to interest rates; therefore, the only way that monetary policy can affect investment spending is through its effects on interest rates." Is this statement true, false, or uncertain? Explain your answer.
9. From March 2009 to February 2020, the S&P 500 stock index increased by over 350%. Over the same period, University of Michigan's index of consumer confidence increased from 57.3 to 101.0 (an increase of 75%). Explain how this relates to the monetary transmission mechanisms.
10. From mid-2008 to early 2009, the Dow Jones Industrial Average declined by more than 50%, while real interest rates were low or falling. What does this scenario suggest should have happened to investment?
11. Nobel Prize winner Franco Modigliani found that the most important transmission mechanisms of monetary policy involve consumer expenditure. Describe how at least two of these mechanisms work.
12. In the late 1990s, the stock market was rising rapidly, the economy was growing, and the Federal Reserve kept interest rates relatively low. Comment on how this policy stance would affect the economy as it relates to the Tobin *q* transmission mechanisms.
13. During and after the global financial crisis, the Bank of Canada reduced the overnight interest rate to nearly zero. At the same time, the stock market fell dramatically and housing market values declined sharply. Comment on the effectiveness of monetary policy during this period with regard to the wealth channel.
14. From August 2014 to August 2017, the Fed continued to reiterate that monetary policy was "accommodative." And during this time, excess reserves by banks decreased from around $2.67 trillion to around $2 trillion, a decline of about 25%. What does this say about the bank lending channel?
15. Why does the credit view imply that monetary policy has a greater effect on small businesses than on large firms?
16. Why might the bank lending channel be less effective today than it once was?
17. If adverse selection and moral hazard increase, how does this affect the ability of monetary policy to address economic downturns?
18. How does the Great Depression demonstrate the unanticipated price level channel?
19. How are the wealth effect and the household liquidity effect similar? How are they different?
20. After the coronavirus pandemic hit the economy, mortgage rates reached record-low levels in 2020.
    a. What effect should this have had on the economy, according to the household liquidity effect channel?
    b. During much of this time, most banks raised their credit standards significantly, making it much more difficult to qualify for home loans and to refinance existing loans. How does this information alter your answer to part (a)?
21. "If the overnight interest rate is at zero, the Bank of Canada can no longer implement effective accommodative policy." Is this statement true, false, or uncertain? Explain.
22. In December 2015, the Fed raised the fed funds rate for the first time in nearly a decade and gradually increased interest rates over a prolonged period of time after that. However, the Fed continued to reiterate in its policy statements during this time that "the stance of monetary policy remains accommodative." Explain this seeming contradiction.
23. In general, if stock prices are rising, consumption growth is strong, house price appreciation is high, and unemployment is low, would you classify monetary policy as likely to be tight or easy?
24. How does the experience of Japan during the "two lost decades" lend support to the four lessons for monetary policy outlined in this chapter?

APPLIED PROBLEMS

25. Suppose the economy is in recession and the monetary policymakers lower interest rates in an effort to stabilize the economy. Use an aggregate supply and demand diagram to demonstrate the effects of a monetary easing when the transmission mechanisms are functioning normally and when the transmission mechanisms are weak, such as during a deep downturn or when significant financial frictions are present.

DATA ANALYSIS PROBLEMS

1. Real-Time Data Analysis A "rate cycle" is a period of monetary policy during which the federal funds rate moves from its low point toward its high point, or vice versa, in response to business cycle conditions. Go to the St. Louis Federal Reserve FRED database and find data on the federal funds rate (FEDFUNDS), real business fixed investment (PNFIC1), real residential investment (PRFIC1), and consumer durable expenditures (PCDGCC96) for the United States. Use the *frequency* setting to convert the federal funds rate data to "Quarterly," and download the data.

a. When did the last rate cycle begin and end? (*Note*: If a rate cycle is currently in progress, use the current period as the end.) Is this rate cycle a contractionary or an expansionary rate cycle?
b. Calculate the percentage change in business fixed investment, residential (housing) investment, and consumer durable expenditures over this rate cycle.
c. Based on your answers to parts (a) and (b), how effective was the traditional interest-rate channel of monetary policy over this rate cycle?

2. Real-Time Data Analysis As defined in Exercise 1, a "rate cycle" is a period of monetary policy during which the federal funds rate moves from its low point toward its high point, or vice versa, in response to business cycle conditions. Go to the St. Louis Federal Reserve FRED database and find data on the federal funds rate (FEDFUNDS), bank reserves (TOTRESNS), bank deposits (TCDSL), commercial and industrial loans (BUSLOANS), real estate loans (REALLN), real business fixed investment (PNFIC96), and real residential investment (PRFIC1). Use the *frequency* setting to convert the federal funds rate, bank reserves, bank deposits, commercial and industrial loans, and real estate loans data to "Quarterly," and download the data.

a. When did the last rate cycle begin and end? (*Note*: If a rate cycle is currently in progress, use the current period as the end.) Is this rate cycle a contractionary or an expansionary rate cycle?
b. Calculate the percentage change in bank deposits, bank lending, real business fixed investment, and real residential (housing) investment over this rate cycle.
c. Based on your answers to parts (a) and (b), how effective was the bank lending channel of monetary policy over this rate cycle?

Glossary

activists Economists who regard the self-correcting mechanism, which works through wage and price adjustment, as very slow because wages and prices are sticky. Activists see the need to pursue active policy to eliminate high unemployment when it develops.

adaptive expectations Expectations for the value of a variable that are based on an average of past values of the variable.

AD/AS diagram A graph that shows the aggregate demand curve, the short-run aggregate supply curve, and the macroeconomic equilibrium.

advances See *overdraft loans*.

adverse selection The problem created by asymmetric information prior to a financial transaction: It occurs when one party to a transaction has information about a hidden characteristic and takes economic advantage of this information by making an agreement (transaction) with less informed parties.

agency theory The analysis of how asymmetric information problems affect economic behaviour.

aggregate demand The total quantity of output demanded in an economy at different price levels.

aggregate demand curve The curve showing the relationship between the price level and the quantity of aggregate output demanded when the goods and money markets are in equilibrium.

aggregate income The total income of factors of production (land, labour, capital) in an economy.

aggregate output The total production of final goods and services in an economy.

aggregate output index An index in which potential GDP is designated as 100 and all other values of real GDP are set relative to 100.

aggregate price level The average price of goods and services in an economy.

aggregate quantity demanded The total quantity of goods and services that households, businesses, governments, and foreigners want to buy (demand) at any given inflation rate during a specific period, such as a year.

aggregate supply curve The curve showing the relationship between the quantity of output supplied and the price level.

American option An option that can be exercised at any time up to the expiration date of the contract.

anchor currency A currency to which other countries' currencies are pegged.

"animal spirits" Waves of optimism and pessimism that affect consumers' and businesses' willingness to spend.

annuities Insurance products that provide a fixed stream of payments.

appreciation Increase in a currency's value.

arbitrage Elimination of a riskless profit opportunity in a market.

asset A financial claim or piece of property that is a store of value.

asset management The acquisition of assets that have a low rate of default and the diversification of asset holdings to increase profits.

asset market approach An approach to determining asset prices that makes use of stocks of assets rather than flows.

asset-price bubbles Occur when asset prices in the stock and real estate markets are driven well above the assets' fundamental economic values by investor psychology.

asset transformation The process of turning risky assets into safer assets, accomplished by creating and selling assets with risk characteristics that people are comfortable with and then using the funds acquired by selling these assets to purchase other assets that have far more risk.

asymmetric information The unequal knowledge that each party to a transaction has about the other party.

automated teller machine (ATM) An electronic machine that provides banking services 24 hours a day.

autonomous consumption expenditure The amount of consumer expenditure that is independent of disposable income.

autonomous easing of monetary policy The lowering of real interest rates at any given inflation rate.

autonomous investment A component of planned investment spending that is completely exogenous and so is unexplained by the variables in a model.

autonomous net exports The level of net exports that is treated as exogenous (outside the model).

autonomous spending Exogenous spending that is unrelated to variables in the model such as output or the real interest rate.

autonomous tightening of monetary policy The raising of real interest rates at any given inflation rate.

balance of payments A bookkeeping system for recording all payments that have a direct bearing on the movement of funds between a country and other countries.

balance sheet A balanced list of the assets and liabilities of a bank (or firm): Total assets equal total liabilities plus capital.

Bank Act reform The legislation that took effect in October 2001 regarding the policy framework of the Canadian financial services sector.

bank failure A situation in which a bank cannot satisfy its obligations to pay its depositors and other creditors and so goes out of business.

bank panic The simultaneous failure of many banks, as during a financial crisis.

bank rate The interest rate the Bank of Canada charges to members of the Canadian Payments Association.

banker's risk The risk of not holding enough reserves to make immediate and larger-than-normal cash payments to liability holders.

banks Financial institutions (such as commercial banks, savings and loan associations, and credit unions) that accept money deposits and make loans.

base money The sum of the Bank of Canada's monetary liabilities (notes outstanding and bank settlement balances) and coins outstanding. Also called *monetary base*.

Basel Accord An agreement that required that banks hold as capital at least 8% of their risk-weighted assets.

Basel Committee on Banking Supervision An international committee of bank supervisors that meets under the auspices of the Bank for International Settlements in Basel, Switzerland.

behavioural finance A subfield of finance that applies concepts from other social sciences, such as anthropology, sociology, and particularly psychology, to explain the behaviour of securities prices.

Board of Directors A board with 15 members (including the governor) that is responsible for the management of the Bank of Canada.

Board of Governors of the Federal Reserve System A board with seven governors (including the chair) that plays an essential role in decision making within the Federal Reserve System.

bond A debt security that promises to make periodic payments to the holder for a specified period of time.

boom See *expansion*.

borrowed reserves A bank's borrowings from the Bank of Canada.

branches A main bank's additional offices in which banking operations are conducted.

Bretton Woods system The international monetary system, in use from 1945 to 1971, in which exchange rates were fixed and the U.S. dollar was freely convertible into gold (by foreign governments and central banks only).

brokerage firms Agents for investors who match buyers with sellers.

brokers Agents for investors; brokers match buyers with sellers.

bubble A situation in which the price of an asset differs from its fundamental market value.

budget deficit Government expenditures in excess of tax revenues.

budget surplus Tax revenues in excess of government expenditures.

business cycles The upward and downward movement of aggregate economic activity.

call option An option contract that provides the right to buy a security at a specified price.

call premium Price of a call option.

call (redemption) The return of money invested when the security matures.

Canadas Bonds issued by the Government of Canada.

capital Wealth, either financial or physical in form, that is employed to produce more wealth.

capital adequacy management A bank's decisions about the amount of capital it should maintain and the best ways to acquire the needed capital.

capital controls Restrictions on the free movement of capital across borders.

capital market A financial market in which longer-term debt (generally with an original maturity of greater than one year) and equity instruments are traded.

capital mobility A situation in which foreigners can easily purchase a country's assets, and the country's residents can easily purchase foreign assets.

cash flows Cash payments to the holder of a security.

central bank The government agency that oversees a country's banking system and is responsible for the amount of money and credit supplied in the economy; in Canada, the Bank of Canada.

certificate of deposit (CD) A financial instrument offered by banks and near banks that requires the customer to leave a deposit untouched for a predetermined period of time.

closed-end fund A mutual fund that sells a fixed number of shares of stock and does not continue to accept investments.

coinsurance A situation in which only a portion of losses is covered by insurance so that the insured suffers a percentage of the losses along with the insurance agency.

collateral Property that is pledged to the lender to guarantee payment in the event that the borrower is unable to make debt payments.

commodity money Money made up of precious metals or another valuable commodity.

common stock A security that is a claim on the earnings and assets of a company.

community banks Small banks with local roots.

compensating balance The required minimum amount of funds that a firm receiving a loan must keep in a chequing account at the lending bank.

complete crowding out The decline in investment that results when an expansionary fiscal policy increases the interest rate.

conflicts of interest A manifestation of the moral hazard problem. Conflicts of interest occur when a financial institution provides multiple services with conflicting goals, a situation that may lead the firm to conceal important information or disseminate misleading information.

consol A perpetual bond with no maturity date and no repayment of principal and that makes periodic, fixed coupon payments.

constant-money-growth-rate rule Monetary policy rule, advocated by monetarists, in which policymakers keep the money supply growing at a constant rate, regardless of the state of the economy.

constrained discretion A framework in which the general objectives and tactics of policymakers—but not specific actions—are committed to in advance; this framework imposes a conceptual structure and inherent discipline on policymakers without eliminating all flexibility.

consumer durable expenditure Spending by consumers on durable items such as automobiles and household appliances.

consumption Spending by consumers on nondurable goods and services (including services related to the ownership of homes and consumer durables).

consumption expenditure The total demand for consumer goods and services.

consumption function The relationship between disposable income and consumer expenditure.

conventional monetary policy tools During normal times the Bank of Canada uses the three tools of monetary policy—open market operations, settlement balances management, and Bank of Canada lending—to control the money supply and interest rates. These three tools are referred to as conventional monetary policy.

cost-push inflation Inflation that occurs when workers push to obtain higher wages.

cost-push shocks Price shocks that occur when workers push for wages that are higher than productivity gains, thereby driving up costs and inflation.

costly state verification Monitoring of a firm's activities, an expensive process in terms of both time and money.

coupon bond A credit market instrument that pays the owner a fixed interest payment every year until the maturity date, at which time a specified final amount is repaid.

coupon rate The dollar amount of the yearly coupon payment expressed as a percentage of the face value of a coupon bond.

credibility A commitment to policy actions that is believed by the public.

credit boom A lending spree in which financial institutions expand their lending at a rapid pace.

credit default swaps Financial insurance contracts that provide payments to holders of bonds if they default.

credit derivatives Derivatives that have payoffs to previously issued securities but ones that bear credit risk.

credit easing Altering the composition of the central bank's balance sheet in order to improve the functioning of particular segments of the credit markets.

credit-linked note A type of credit derivative that combines a bond and a credit option.

credit options Options in which, for a fee, the purchaser has the right to get profits that are tied either to the price of an underlying risky security or to an interest rate.

credit rationing Occurs when a lender refuses to make loans even though borrowers are willing to pay the stated interest rate or even a higher rate, or when a lender restricts the sizes of loans so that the loans are made for less than the full amount sought.

credit risk The risk arising from the possibility that the borrower will default.

credit spread The difference between the interest rate on loans to households and businesses and the interest rate on completely safe assets that are sure to be paid off, such as U.S. Treasury securities.

credit swap A transaction in which risky payments on loans are swapped for each other.

credit view Monetary transmission mechanisms that operate through asymmetric information effects on credit markets.

credit-rating agencies Investment advisory firms that rate the quality of corporate and municipal bonds in terms of the probability of default.

currency Paper money (such as dollar bills) and coins.

currency board A monetary regime in which the domestic currency is backed 100% by a foreign currency (say, dollars) and in which the note-issuing authority, whether it be the central bank or the government, establishes a fixed exchange rate to this foreign currency and stands ready to exchange domestic currency at this rate whenever the public so requests.

currency mismatch Refers to how exchange rate changes affect the present value of income and expenditure flows.

currency swaps The exchange of a set of payments in one currency for a set of payments in another currency.

current account An account that shows international transactions involving currently produced goods and services.

current account balance The sum of trade balance, net interest income, and transfers.

current yield An approximation of the yield to maturity; equal to the yearly coupon payment divided by the price of a coupon bond.

data lag The time it takes for policymakers to obtain data on the state of the economy.

dealers People who link buyers with sellers by buying and selling securities at stated prices.

debt deflation A situation in which a substantial decline in the price level sets in, leading to a further deterioration in firms' net worth because of the increased burden of indebtedness.

deductible An amount of any loss that must be paid by the insured before the insurance company will pay anything.

default A situation in which the party issuing a debt instrument is unable to make interest payments or pay off the amount owed when the instrument matures.

default-free bonds Bonds with no default risk, such as U.S. government bonds.

defined-benefit plan A pension plan in which the benefits are stated up front and are paid regardless of how the investments perform.

defined contribution plan A pension plan in which the contributions are stated up front but the benefits paid depend on the performance of the investments.

deleveraging Occurs when financial institutions cut back on their lending because they have less capital.

demand curve The curve depicting the relationship between quantity demanded and price when all other economic variables are held constant.

demand for money The quantity of money that people want to hold.

demand shocks Shocks that can shift the aggregate demand curve, including changes in the money supply, changes in government expenditure and taxes, changes in net exports, changes in consumer and business spending, and financial frictions.

demand-pull inflation Inflation that results when policymakers pursue policies that shift the aggregate demand curve.

demutualization Refers to the conversion of a mutual company to a share company.

deposit facility The European Central Bank's standing facility, in which banks are paid a fixed interest rate 100 basis points below the target financing rate.

deposit outflows Losses of deposits when depositors make withdrawals or demand payment.

deposit rate ceiling A restriction on the maximum interest rate payable on deposits.

depreciation A decrease in a currency's value.

desired reserve ratio The fraction of deposits that banks desire to keep as reserves.

desired reserves Reserves that are held to meet the banks' desire that for every dollar of deposits, a certain fraction should be kept as reserves.

devaluation The resetting of the fixed value of a currency to a lower level.

direct clearers Members of the Canadian Payments Association who participate directly in the Automated Clearing Settlement System (ACSS) and maintain a settlement account at the Bank of Canada.

discount bond A credit market instrument that is bought at a price below its face value and whose face value is repaid at the maturity date; it does not make any interest payments. Also called a *zero-coupon bond*.

discount rate The interest rate that the Bank of Canada charges banks on discount loans.

discount window The Federal Reserve facility at which discount loans are made to banks.

discretion When policymakers make decisions that they believe are the best policies to address a given situation at a given point in time.

disintermediation A reduction in the flow of funds into the banking system that causes the amount of financial intermediation to decline.

disposable income Total income available for spending, equal to aggregate income minus taxes.

diversification Investing in a collection (portfolio) of assets whose returns do not always move together, with the result that the overall risk is lower than that on the individual assets.

dividends Periodic payments made by equities to shareholders.

divine coincidence A phrase coined by Olivier Blanchard, referring to the situation in which a policy meant to stabilize inflation is also the best policy for stabilizing economic activity.

dollarization The adoption of a sound currency, like the U.S. dollar, as a country's money.

dominated assets Assets such as currency and chequable assets, which earn lower returns than other assets that are just as safe.

drawdowns Transfers of government deposits from the government's accounts with the direct clearers to the government's account with the Bank of Canada.

dual banking system The banking system in place in the United States, in which banks supervised by the federal government and banks supervised by the states operate side by side.

dual mandate A central bank mandate that features two co-equal objectives: price stability and maximum employment.

duration How many years an investor has to wait until they receive the asset's discounted present value.

duration analysis A measurement of the sensitivity of the market value of a bank's assets and liabilities to changes in interest rates.

duration gap analysis A measurement of the sensitivity of the market value of a bank's assets and liabilities to changes in interest rates.

e-cash Electronic money that is used on the Internet to purchase goods or services.

e-finance A new means of delivering financial services electronically.

economies of scale The reduction in transaction costs per dollar of transaction as the size (scale) of transactions increases.

economies of scope The ability to use one resource to provide many different products and services.

effectiveness lag The time it takes for a policy to actually have an impact on the economy.

efficient market hypothesis The application of the theory of rational expectations to financial markets. Also called the *theory of efficient capital markets*.

electoral business cycles Persistent cyclical patterns of key policy and target variables across electoral terms.

electronic money (e-money) Money that exists only in electronic form and substitutes for cash as well.

emerging market countries Countries that have recently opened up to flows of goods, services, and capital from the rest of the world.

equation of exchange The equation $MV = PY$, which relates nominal income to the quantity of money.

equities Claims to share in the net income and assets of a corporation (such as common stock).

equity capital See *net worth*.

equity multiplier (EM) The amount of assets per dollar of equity capital.

Eurobonds Bonds denominated in a currency other than that of the country in which they are sold.

Eurocurrencies A variant of the Eurobond; foreign currencies deposited in banks outside the home country.

Eurodollars U.S. dollars that are deposited in foreign banks outside the United States or in foreign branches of U.S. banks.

European option An option that can be exercised only at the expiration date of the contract.
Eurosystem Refers to the European Central Bank and the national central banks of the countries that have adopted the euro.
excess demand A situation in which the quantity demanded is greater than the quantity supplied.
excess reserves Reserves in excess of desired reserves.
excess supply A situation in which the quantity supplied is greater than the quantity demanded.
Exchange Fund Account The fund that holds Canada's official foreign exchange assets.
exchange rate The price of one currency stated in terms of another currency.
exchange-rate peg The fixing of the value of the domestic currency to the value of another currency so that the exchange rate is fixed. Also called *exchange-rate targeting*.
exchange-rate targeting See *exchange-rate peg*.
exchanges Secondary markets in which buyers and sellers of securities (or their agents or brokers) meet in one central location to conduct trades.
exercise price (strike price) The price at which the purchaser of an option has the right to buy or sell the underlying financial instrument.
exogenous Independent of variables in the model.
expansion A period when aggregate economic activity is increasing or expanding (also called a *boom*).
expectations theory The proposition that the interest rate on a long-term bond will equal the average of the short-term interest rates that people expect to occur over the life of the long-term bond.
expected return The return on an asset expected over the next period.
face value A specified final amount paid to the owner of a coupon bond at the maturity date. Also called *par value*.
federal funds rate The interest rate on overnight loans of deposits at the Federal Reserve.
Federal Open Market Committee (FOMC) The committee that makes decisions regarding the conduct of open market operations; composed of the seven members of the Board of Governors of the Federal Reserve System, the president of the Federal Reserve Bank of New York, and the presidents of four other Federal Reserve banks (on a rotating basis).
Federal Reserve Banks The 12 district banks in the Federal Reserve System.
Federal Reserve System (the Fed) The central banking authority responsible for monetary policy in the United States.
fiat money Paper currency decreed by a government as legal tender but not convertible into coins or precious metal.
finance paper A money market security issued by a corporation to obtain funds to meet short-term obligations.
financial account Account that shows the international transactions that involve the purchase or sale of assets.
financial account balance The difference between the net acquisition of financial assets and the net incurrence of financial liabilities.
financial crisis A major disruption in financial markets that is characterized by sharp declines in asset prices and the failures of many financial and nonfinancial firms.
financial derivatives Instruments that have payoffs that are linked to previously issued securities; used as risk reduction tools.
financial engineering The process of researching and developing new financial products and services that will meet customer needs and also prove profitable.
financial frictions Asymmetric information problems that act as barriers to the efficient allocation of capital.
financial futures contract A futures contract to buy or sell a financial instrument.
financial futures option An option in which the underlying instrument is a futures contract. Also called a *futures option*.
financial globalization The process of opening up to flows of capital and financial firms from other nations.
financial innovation The introduction of new types of financial products in an economy.
financial intermediaries Institutions (such as banks, insurance companies, mutual funds, pension funds, and finance companies) that borrow funds from people who have saved and then make loans to others.
financial intermediation The process of indirect finance whereby financial intermediaries link lender-savers with borrower-spenders.
financial liberalization The elimination of restrictions on financial markets.
financial markets Markets in which funds are transferred from people who have a surplus of available funds to people who have a shortage of available funds.
financial panic The widespread collapse of financial markets and intermediaries in an economy.
financial supervision (prudential supervision) The oversight of financial institutions and their operations.
fire sales Forced, rapid sales of assets to raise needed funds.
fiscal policy Policy that involves decisions about government spending and taxation.
Fisher effect Describes the rise in interest rates that occurs when inflation is at the level of expected inflation; named after economist Irving Fisher.
fixed exchange rate regime A regime in which central banks buy and sell their own currencies to keep their exchange rates fixed at a certain level.
fixed investment Spending by firms on equipment (computers, airplanes) and structures (factories, office buildings), and planned spending on residential housing.
fixed-payment loan A credit market instrument that provides the borrower with an amount of money that is repaid through fixed periodic (usually monthly) payments made over a set number of years.

float Cash items in process of collection at the Bank of Canada minus deferred-availability cash items.
floating exchange rate regime An exchange rate regime in which the values of currencies are allowed to fluctuate against one another.
foreign bonds Bonds sold in a foreign country and denominated in that country's currency.
foreign exchange intervention An international financial transaction in which a central bank buys or sells currency to influence foreign exchange rates.
foreign exchange market The market in which exchange rates are determined.
foreign exchange rate See *exchange rate*.
forward contracts An agreement by two parties to engage in a financial transaction at a future (forward) point in time.
forward exchange rate The exchange rate for a forward transaction.
forward guidance A central bank commitment to a future path of the policy interest rate.
forward transaction A transaction that involves the exchange of bank deposits denominated in different currencies at some specified future date.
four-pillar approach Regulation of the banking industry by institution (banking, brokerage, trusts, and insurance) versus regulation by function.
free banking A system that permits the organization of a bank by any group that meets certain established criteria concerning the amount of equity capital and maintenance of reserves.
free-rider problem The problem that occurs when people who do not pay for information take advantage of the information that other people have paid for.
fully funded Describing a pension plan in which the contributions to the plan and their earnings over the years are sufficient to pay out the defined benefits when they come due.
fundamental economic values Values of assets based on realistic expectations of their future income streams.
futures contract A contract in which the seller agrees to provide a certain standardized commodity to the buyer on a specific future date at an agreed-on price.
gap analysis A measurement of the sensitivity of bank profits to changes in interest rates, calculated by subtracting the amount of rate-sensitive liabilities from the amount of rate-sensitive assets.
generalized dividend model A model in which the price of a stock is determined only by the present value of the dividends.
Glass-Steagall An Act passed in 1933 by the U.S. Congress to legally separate investment banking from commercial banking.
goal independence The ability of the central bank to set the goals of monetary policy.
gold standard A fixed exchange rate regime under which a currency is directly convertible into gold.
Gordon growth model A simplified model used to compute the value of a stock by assuming constant dividend growth.
Governing Council A council with six members (including the governor) that is responsible for the management of the Bank of Canada.
government budget constraint The requirement that the government budget deficit must equal the sum of the change in the monetary base and the change in government bonds held by the public.
government purchases Spending by all levels of government (federal, provincial, and local) on goods and services.
government-sponsored enterprises Sponsored agencies that function as private corporations but have close ties to the government.
gross domestic product (GDP) The value of all final goods and services produced in the economy during the course of a year.
haircuts The excess amount of collateral above the amount of a loan.
hedge To protect oneself against risk.
hedge fund A special type of mutual fund that acquires funds by selling shares but only to very wealthy people, so they are less regulated than mutual funds.
hierarchical mandate A central bank mandate in which the goal of price stability comes first; once price stability has been achieved, other goals can be pursued.
hyperinflation An extreme inflation in which the inflation rate exceeds 50% per month.
implementation lag The time it takes for policymakers to change policy instruments once they have decided on a new policy.
impossible trinity Another term for the policy trilemma, in which a country cannot pursue the following three policies at the same time: free capital mobility, a fixed exchange rate, and an independent monetary policy.
incentive-compatible Used to describe a contract in which the incentives of both parties to the contract are in alignment.
income The flow of earnings.
indebtedness In describing a country, the total amount it has borrowed from banks.
inflation The economic condition of a continually rising price level.
inflation gap The difference between inflation and the inflation target.
inflation hedges Alternative assets whose real returns are less affected than that of money when inflation varies.
inflation rate The rate of change of the price level, usually measured as a percentage change per year.
inflation shocks Shifts in inflation that are independent of the amount of slack in the economy or expected inflation.
inflation target A central bank target for the inflation rate.
inflation targeting A monetary policy strategy that involves the public announcement of a medium-term numerical target for inflation.
initial public offering (IPO) A corporation's first sale of securities to the public.

instrument independence The ability of the central bank to set monetary policy instruments.
interest rate The cost of borrowing or the price paid for the rental of funds (usually expressed as a percentage per year).
interest-rate forward contracts A forward contract that is linked to a debt instrument.
interest-rate risk The possible reduction in returns associated with changes in interest rates.
interest-rate swaps A financial contract that allows one party to exchange (swap) a set of interest payments for another set of interest payments owned by another party.
intermediate target Any of a number of variables, such as monetary aggregates or interest rates, that have a direct effect on employment and the price level and that the Bank of Canada seeks to influence.
intermediate term With reference to a debt instrument, having a maturity of between 1 and 10 years.
International Monetary Fund (IMF) The international organization created by the Bretton Woods agreement whose objective it is to promote the growth of world trade by making loans to countries experiencing balance-of-payments difficulties.
international reserves Central bank holdings of assets denominated in foreign currencies.
inventory investment Spending by firms on additional holdings of raw materials, parts, and finished goods.
inverted yield curve A yield curve that is downward-sloping.
investment banks Firms that assist in the initial sale of securities in the primary market.
***IS* curve** The relationship that describes the combinations of aggregate output and interest rates for which the total quantity of goods produced equals the total quantity demanded (goods market equilibrium).
junk bonds Bonds with ratings below Baa (or BBB) that have a high default risk.
Keynesian A follower of John Maynard Keynes who believes that movements in the price level and aggregate output are driven by changes not only in the money supply but also in government spending and fiscal policy, and who does not regard the economy as inherently stable.
Large Value Transfer System (LVTS) An electronic net settlement system for the transfer of large-value payments.
legislative lag The time it takes to pass legislation to implement a particular policy.
lender of last resort A lender that provides reserves to financial institutions when no one else is willing to do so; such lending is usually done to prevent a financial crisis.
leverage cycle A feedback loop in which a boom in the issuance of credit leads to higher asset prices, which results in higher capital buffers at financial institutions, which supports further lending, which raises prices further, and so on; during the subsequent bust, asset prices fall, leading to cuts in lending, declines in asset prices, and so on.
leverage ratio A bank's capital divided by its assets.
liabilities IOUs or debts.
liability management The acquisition of funds at low cost to increase profits.
liquid Easily converted into cash.
liquidity The relative ease and speed with which an asset can be converted into cash.
liquidity management The decisions made by a bank in order to maintain enough liquid assets to meet the bank's obligations to depositors.
liquidity preference framework A model developed by John Maynard Keynes that predicts the equilibrium interest rate on the basis of the supply of and demand for money.
liquidity preference theory John Maynard Keynes's theory of the demand for money.
liquidity premium theory The theory that the interest rate on a long-term bond will equal an average of the short-term interest rates expected to occur over the life of the long-term bond, plus a positive term (liquidity) premium.
liquidity services Services that financial intermediaries provide to their customers to make it easier for the customers to conduct their transactions.
liquidity trap A case of ultrasensitivity of the demand for money to interest rates, in which conventional monetary policy has no direct effect on aggregate spending because a change in the money supply has no effect on the interest rate.
***LM* curve** The positive relationship between the interest rate and real output that arises in the market for real money balances while holding the price level fixed.
load fund A mutual fund that charges a fee when money is added to or withdrawn from the fund.
loan commitment A bank's commitment to provide a firm with loans (for a specified period of time) up to a given amount at an interest rate that is tied to some market interest rate.
loan sale The sale through a contract (also called a *secondary loan participation*) of all or part of the cash stream from a specific loan, thereby removing the loan from the bank's balance sheet.
long in a call Purchase of a call option.
long position A contractual obligation to take delivery of an underlying financial instrument.
long-run monetary neutrality The property that in the long run a change in the supply of nominal money balances by the central bank does not influence real variables.
long term With reference to a debt instrument, having a maturity of 10 years or more.
loanable funds The funds the savers in the economy decided to save and lend out to borrowers.
M1 A measure of money that includes currency, traveller's cheques, and chequable deposits.
M1+ Includes currency, personal chequing accounts, and current accounts, plus other assets that have cheque-writing features—all chequable notice deposits at chartered banks, TMLs, and CUCPs.
M1++ M1, plus all nonchequable notice deposits at chartered banks, TMLs, and CUCPs.

M2 A measure of money that includes M1 plus the following: money market deposit accounts, money market mutual fund shares, small-denomination time deposits, savings deposits, overnight repurchase agreements, and overnight Eurodollars.

M2+ Includes M2 plus deposits at near banks, life insurance company annuities, and money market mutual funds.

M2++ Adds to M2 Canada Savings Bonds and non-money market mutual funds.

M3 Adds to M2 non-personal term deposits at chartered banks and foreign currency deposits of residents at chartered banks.

macroeconometric models Models used to forecast the effects of policy on economic activity; macroeconometric models use collections of equations that describe statistical relationships among many economic variables.

macroeconomic equilibrium Occurs when all markets are simultaneously in equilibrium, at the point where the quantity of aggregate output demanded equals the quantity of aggregate output supplied.

macro hedge A hedge of interest-rate risk for a financial institution's entire portfolio.

macroprudential regulation Regulatory policy that affects the credit markets in the aggregate.

macroprudential supervision Supervision that focuses on the safety and soundness of the financial system in the aggregate.

main refinancing operations Weekly reverse transactions (purchases or sales of eligible assets through repurchase agreements or credit operations, with eligible assets as collateral) that are reversed within two weeks and are the primary monetary policy tool of the European Central Bank.

managed float regime An exchange rate regime in which countries attempt to influence their exchange rates by buying and selling currencies (also called a *dirty float*).

management of expectations A term coined by Michael Woodford that refers to a commitment by the Fed to keep the federal funds rate at zero for an extended period in order to generate a fall in the long-term interest rate.

margin requirement A sum of money that must be kept in an account (the margin account) at a brokerage firm.

marginal lending facility The European Central Bank's standing lending facility, in which banks can borrow (against eligible collateral) overnight loans from the national central bank at a rate that is 100 basis points above the target financing rate.

marginal lending rate The interest rate charged by the European Central Bank for borrowing at its marginal lending facility.

marginal propensity to consume The slope of the consumption function line that measures the change in consumer expenditure that results from an additional dollar of disposable income.

marked to market Repriced and settled in the margin account at the end of every trading day to reflect any change in the value of the future contract.

market equilibrium Occurs in an economy when the quantity that people are willing to buy (demand) is equal to the quantity that people are willing to sell (supply).

market fundamentals Items that have a direct impact on the future income streams of a security.

maturity The time to the expiration date (maturity date) of a debt instrument.

medium of exchange Anything that is used to pay for goods and services.

micro hedge A hedge for a specific asset.

microprudential supervision Supervision that focuses on the safety and soundness of individual financial institutions.

monetarism The doctrine that changes in the money supply are the primary cause of business cycles.

monetarist A follower of Milton Friedman who sees changes in the money supply as the primary source of movements in the price level and aggregate output and who views the economy as inherently stable.

monetary aggregates The measures of the money supply used by the Bank of Canada.

monetary base The sum of the Bank of Canada's monetary liabilities (notes outstanding and bank settlement balances) and coins outstanding.

monetary conditions Determined by the level of short-term interest rates and the exchange rate of the Canadian dollar.

monetary policy The management of the money supply and interest rates.

monetary policy (*MP*) curve The curve showing the relationship between the real interest rate set by the central bank and the inflation rate.

monetary targeting A monetary policy strategy in which the central bank announces that it will achieve a certain value (the target) of the annual growth rate of a monetary aggregate.

monetary theory The theory that relates changes in the quantity of money to changes in economic activity.

monetary (currency) union A group of countries that decides to adopt a common currency, thereby fixing the countries' exchange rates in relation to each other.

monetizing the debt A method of financing government spending whereby the government debt issued to finance government spending is removed from the hands of the public and is replaced by high-powered money instead. Also called *printing money*.

money (money supply) Anything that is generally accepted as payment for goods or services or in the repayment of debts.

money centre banks Large banks in key financial centres (New York, Chicago, San Francisco).

money market A financial market in which only short-term debt instruments (generally those with original maturities of less than one year) are traded.

money multiplier A ratio that relates the change in the money supply to a given change in the monetary base.

monoline insurance companies Companies that specialize in credit insurance.

moral hazard The risk that occurs when an informed party takes a hidden (unobserved) action that harms the less-informed party.

mortgage-backed securities Securities that cheaply bundle and quantify the default risk of the underlying high-risk mortgages.

mortgages Loans to households or firms that wish to purchase housing, land, or other real structures; the structure or land itself serves as collateral for the loan.

multilateral netting The LVTS process by which only the net credit or debit position of each participant vis-à-vis all other participants is calculated for settlement, thereby reducing the need for a large amount of settlement balances.

multiple deposit creation The process whereby, when the Bank of Canada supplies the banking system with $1 of additional reserves, deposits increase by a multiple of this amount.

municipal bonds (municipals) Bonds issued by municipalities.

national banks Federally chartered banks.

natural rate of output See *potential output*.

natural rate of unemployment The rate of unemployment consistent with full employment; the rate at which the demand for labour equals the supply of labour.

net exports Net foreign spending on domestic goods and services, equal to exports minus imports.

net worth The difference between a firm's assets (what it owns or is owed) and its liabilities (what it owes). Also called *equity capital*.

no-load fund A mutual fund that does not charge a fee when funds are added to or withdrawn from the fund.

nominal anchor A nominal variable, such as the inflation rate, an exchange rate, or the money supply, that monetary policymakers use to tie down the price level.

nominal GDP targeting A monetary policy strategy in which the central bank announces an objective of hitting a particular level of nominal GDP (real GDP times the price level) that grows over time.

nominal interest rate An interest rate that does not take inflation into account.

nonaccelerating inflation rate of unemployment (NAIRU) The rate of unemployment at which there is no tendency for inflation to change.

nonactivists Economists who believe that wages and prices are very flexible and so the self-correcting mechanism works very rapidly; nonactivists do not see the need to pursue policies to return the economy to full employment.

nonborrowed monetary base The monetary base minus discount loans (borrowed reserves).

nonconventional monetary policy tools The non-interest-rate tools used by central banks to stimulate the economy: liquidity provision, asset purchases, and commitments to future monetary policy actions.

nontradable Goods and services that are not traded across borders.

notional principal The amount on which interest is being paid in a swap arrangement.

off-balance-sheet activities Bank activities that involve the trading of financial instruments and the generation of income from fees and loan sales, all of which affect bank profits but are not visible on bank balance sheets.

open-end fund A mutual fund that accepts investments and allows investors to redeem shares at any time. The value of the shares is tied to the value of investment assets of the fund.

open interest The number of contracts outstanding.

open market buyback operations The Bank of Canada's buying or selling of bonds in the open market.

open market operations The Bank of Canada's buying or selling of bonds in the open market.

open market purchase A purchase of bonds by the Bank of Canada.

open market sale A sale of bonds by the Bank of Canada.

operating band The Bank's operational objective is to keep the overnight rate within an operating band of 25 basis points.

operating instrument A variable that is very responsive to the central bank's tools and indicates the stance of monetary policy (also called a *policy instrument*).

opportunity cost The amount of interest (expected return) sacrificed by not holding an alternative asset.

optimal forecast The best guess of future conditions, made using all available information.

option A contract that gives the purchaser the option (right) to buy or sell the underlying financial instrument at a specified price, called the exercise price or strike price, within a specific period of time (the term to expiration).

originate-to-distribute model A business model in which a mortgage is originated by a separate party, typically a mortgage broker, and then distributed to an investor as an underlying asset in a security.

output gap The percentage difference between aggregate output and potential output.

over-the-counter (OTC) market A secondary market in which dealers at different locations who have an inventory of securities stand ready to buy and sell securities "over the counter" to anyone who comes to them and is willing to accept their prices.

overdraft loans Borrowings from the Bank of Canada. Also called *advances*.

overnight cash rate The interest rate for very-short-term interbank loans in the euro area.

overnight interest rate The interest rate at which participants borrow and lend overnight funds to each other in the money market.

overnight rate The interest rate on overnight (one-day) securities.

partisan business cycles Systematic and permanent differences in macroeconomic outcomes that differ by political party.

payment technology Methods of payment that include credit cards and electronic payments.

payments system The method of conducting transactions in an economy.

perpetuity See *consol*.

personal pension plans Pension plans designed for business owners, associations, and professionals.

Phillips curve theory A theory suggesting that changes in inflation are influenced by the state of the economy relative to its production capacity as well as other factors.

planned expenditure The total amount that households, businesses, the government, and foreigners want to spend on domestically produced goods and services.

planned investment spending Total planned spending by businesses on new physical capital (e.g., machines, computers, apartment buildings) plus planned spending on new homes.

policy instrument A variable that is very responsive to the central bank's tools and that indicates the stance of monetary policy (also called an *operating instrument*).

policy rate The target for the Bank of Canada's overnight interest rate.

policy trilemma The idea that a country cannot pursue the following three policies at the same time: free capital mobility, a fixed exchange rate, and an independent monetary policy.

political business cycle A business cycle caused by expansionary policies that are pursued prior to an election.

portfolio A collection or group of assets.

potential GDP See *potential output*.

potential output The level of aggregate output produced when the economy is at full employment (the unemployment rate is at the natural rate of unemployment) and firms are on average producing at their long-run, profit-maximizing level (also called the *natural rate of output*).

preferred habitat theory A theory that holds that the interest rate on a long-term bond is equal to an average of the short-term interest rates expected to occur over the life of the long-term bond plus a positive term premium. Closely related to the liquidity premium theory.

premium The amount paid for an option contract.

present discounted value See *present value*.

present value Today's value of a payment to be received in the future, when the interest rate is i. Also called *present discounted value*.

price stability Low and stable inflation.

primary dealers Government securities dealers, operating out of private firms or commercial banks, with whom the Bank of Canada's open market desk trades.

primary market A financial market in which new issues of a security are sold to initial buyers.

principal–agent problem A moral hazard problem that occurs when the managers in control (the agents) act in their own interest rather than in the interest of the owners (the principals) due to differing sets of incentives.

printing money See *monetizing the debt*.

private-equity firm A financial intermediary that has a similar structure to a venture capital firm but instead of investing in new businesses, it buys the shares of *existing* corporations.

private-equity fund A fund that makes investments in equity securities according to investment strategies associated with private equity.

provincial bonds (provincials) Bonds issued by provinces.

prudential supervision See *financial supervision*.

put option An option contract that provides the right to sell a security at a specified price.

quantitative easing An expansion of the central bank's balance sheet.

quantity theory of money The theory that nominal income is determined solely by movements in the quantity of money.

quotas Restrictions on the quantity of foreign goods that can be imported.

random walk Movements of a variable whose future values cannot be predicted (are random) because, given today's value, the value of the variable is just as likely to fall as it is to rise.

rate of capital gain The change in a security's price relative to its initial purchase price.

rate of return See *return*.

rational expectations Expectations that reflect optimal forecasts (the best guess of future conditions) made using all available information.

real exchange rate The rate at which domestic goods can be exchanged for foreign goods; in other words, the price of domestic goods relative to foreign goods denominated in domestic currency.

real interest rate The interest rate adjusted for expected changes in the price level (inflation) so that it more accurately reflects the true cost of borrowing.

real money balances The quantity of money in real terms.

real terms Terms that reflect the actual amount of goods and services that one can buy.

recession A period during which aggregate output is declining.

recognition lag The time it takes for policymakers to feel confident about the signals the data are sending regarding the future course of the economy.

redeposits Transfers of government deposits from the Bank of Canada to direct clearers.

reference rate The interest rate at which participants borrow and lend overnight funds to each other in the money market.

registered bonds A bond that has its owner registered with the issuer of the bond.

regulatory arbitrage A process in which banks keep on their books assets that are relatively risky, such as a loan to a company with a very low credit rating, while taking off their books low-risk assets, such as a loan to a company with a very high credit rating. In regulatory arbitrage, the risky and low-risk assets have the same risk-based capital requirements.

regulatory forbearance Refraining from exercising a regulatory right to put insolvent institutions out of business.

reinsurance Allocating a portion of the risk to another company in exchange for a portion of the premium.

repos See *special Purchase and Resale Agreements*.

repurchase agreement (repo) An arrangement whereby the Fed or another party purchases securities with the understanding that the seller will repurchase them in a short period of time, usually less than a week.

reserve currency A currency, such as the U.S. dollar, that is used by other countries to denominate the assets they hold as international reserves.

reserves Banks' holding of deposits in accounts with the Bank of Canada plus currency that is physically held by banks (vault cash).

residual claimant A stockholder's right to receive whatever remains after all other claims against a firm's assets have been satisfied.

restrictive covenants Provisions that restrict and specify certain activities that a borrower can engage in.

return The payments to the owner of a security plus the change in the security's value, expressed as a fraction of its purchase price. More precisely called the *rate of return*.

return on assets (ROA) Net profit after taxes per dollar of assets.

return on equity (ROE) Net profit after taxes per dollar of equity capital.

revaluation The resetting of the fixed value of a currency at a higher level.

reverse repos See *Sale and Repurchase Agreements*.

reverses Banks' holdings of deposits in accounts with the central bank plus currency that is physically held by banks (vault cash).

reverse transactions Purchase or sale of eligible assets by the central bank under repurchase agreements or credit operations against eligible assets as collateral that are reversed within two weeks.

risk The degree of uncertainty associated with the return on an asset.

risk premium The spread between the interest rate on bonds with default risk and the interest rate on default-free bonds.

risk sharing The process of creating and selling assets with risk characteristics that people are comfortable with and then using the funds acquired by selling these assets to purchase other assets that have far more risk.

risk structure of interest rates The relationships among the interest rates on various bonds with the same term to maturity.

rules Binding plans that specify how policy will respond (or not respond) to particular data, such as data on unemployment and inflation.

Sale and Repurchase Agreements (SRAs) The Bank of Canada's sale of government securities to primary dealers with an agreement to repurchase them one business day later.

Sarbanes-Oxley Act Passed by U.S. Congress in 2002, Sarbanes-Oxley aims to strengthen the integrity of the auditing process and the quality of information provided in corporate financial statements.

Schedule I banks Comprising the Big Six, together with the Laurentian Bank of Canada, the Canadian Western Bank, and another 14 domestic banks.

Schedule II banks Comprising 55 banks, including a number of domestic Schedule II banks and foreign bank subsidiaries.

Schedule III banks Foreign banks allowed to branch directly into Canada, under certain restrictions.

seasoned issues Securities that have been trading publicly long enough to have let the market clearly establish their value.

secondary market A financial market in which securities that have previously been issued (and are thus secondhand) can be resold.

secondary reserves Short-term Canadian government securities held by banks.

secured debt Debt guaranteed by collateral.

securitization The process of transforming illiquid financial assets into marketable capital market instruments.

security A claim on a borrower's future income that is sold by the borrower to the lender. Also called a *financial instrument*.

segmented markets theory A theory of term structure that sees the markets for different-maturity bonds as completely separated and segmented so that the interest rate on bonds of a given maturity is determined solely by supply of and demand for bonds of that maturity.

seignorage The revenue that a government receives by issuing money.

self-correcting mechanism A characteristic of the economy that causes output to return eventually to the natural rate level, regardless of where it lies initially.

settlement balances Deposits held by directly clearing members of the Canadian Payments Association at the Bank of Canada. They are also known as *clearing balances*.

shadow banking system A system in which bank lending is replaced by lending via the securities market.

short in a call Sale of a call option.

short position A contractual obligation to deliver an underlying financial instrument.

short sales Borrowing stock from brokers and then selling the stock in the market with the hope that a profit will be earned by buying the stock back again ("covering the short") after it has fallen in price.

short term With reference to a debt instrument, having a maturity of one year or less.

simple deposit multiplier The multiple increase in deposits generated by an increase in the banking system's reserves. The behaviour of depositors and banks plays no role in this simple model.

simple loan A credit market instrument that provides the borrower with an amount of funds that must be repaid to the lender at the maturity date, along with an additional payment (interest).

smart card A stored-value card that contains a computer chip that allows the card to be loaded with digital cash from the owner's bank account whenever needed.

sovereign loans Loans to foreign governments and their agencies in the less developed countries.

sovereign wealth fund A state-owned investment fund that invests in foreign assets.

special Purchase and Resale Agreements (special PRAs or SPRAs) The Bank of Canada's purchase of government securities from primary dealers with an agreement to resell them one business day later.

specials See *special Purchase and Resale Agreements*.

speculative attack Involves massive sales of a weak currency or purchases of a strong currency that can cause a sharp change in the exchange rate.

spot exchange rate The exchange rate for a spot transaction.

spot transaction The predominant type of exchange-rate transaction, involving the immediate exchange of bank deposits denominated in different currencies.

stagflation A situation of rising inflation but a falling level of aggregate output.

standing lending facility A lending facility in which healthy banks are allowed to borrow all they want from a central bank.

standing liquidity facilities Refers to the Bank of Canada standing ready to lend to or borrow from a participant to bring their settlement balances to zero at the end of the banking day.

state banks State-chartered banks.

state-owned banks Banks that are owned by governments.

sterilized foreign exchange intervention A foreign exchange intervention that is accompanied by an offsetting open market operation that leaves the monetary base unchanged.

sticky When wages and prices are slow to adjust.

stock A security that is a claim on the earnings and assets of a company.

stock option An option on an individual stock.

stockholders Those who hold stock in a corporation.

store of value A repository of purchasing power over time.

stress tests Tests used by financial institutions to calculate potential losses and capital needs under fictional dire scenarios.

structured credit products Securities that are derived from the cash flows of underlying assets and are tailored to have particular risk characteristics that appeal to investors with different preferences.

subprime mortgages Mortgages made to borrowers with less-than-stellar credit records.

supply curve A curve depicting the relationship between quantity supplied and price when all other economic variables are held constant.

supply shock Any change in technology or the supply of raw materials that can shift the aggregate supply curve.

swap A financial contract that obligates one party to exchange (swap) a set of payments it owns for a set of payments owned by another party.

sweep account An arrangement whereby any balances above a certain amount that remain in a corporation's chequing account at the end of the business day are "swept out" of the account and invested in overnight repos that pay the corporation interest.

systemic risk The risk to the entire payments system due to the inability of one financial institution to fulfill its payment obligations in a timely fashion.

systemically important financial institutions (SIFIs) Firms that are designated by the Financial Stability Oversight Council as systemically important and so are subject to additional oversight and regulation by the Federal Reserve.

T-account A simplified balance sheet with lines in the form of a T that lists only the changes that occur in balance sheet items starting from some initial balance sheet position.

target financing rate The European Central Bank's target for the overnight cash rate, the interest rate for very-short-term interbank loans in the euro area.

tariffs Taxes on imported goods.

Taylor principle The principle that the monetary authorities should raise nominal interest rates by more than the increase in the inflation rate.

Taylor rule Economist John Taylor's monetary policy rule that explains how the overnight interest rate target is set.

term Purchase and Resale Agreements (term PRAs) The Bank of Canada's purchase of government securities from primary dealers with an agreement to resell them at a specified future date.

term repos See *term Purchase and Resale Agreements (PRAs)*.

term structure of interest rates The relationships among interest rates on bonds with different terms to maturity.

theory of efficient capital markets See *efficient market hypothesis*.

theory of portfolio choice A theory that outlines how much of an asset people will want to hold in their portfolios, as determined by wealth, expected returns, risk, and liquidity.

theory of purchasing power parity (PPP) The theory that exchange rates between any two currencies will adjust to reflect changes in the price levels of the two countries.

time-inconsistency problem The problem that occurs when monetary policymakers conduct monetary policy in a discretionary way and pursue expansionary policies that are attractive in the short run but lead to poor long-run outcomes.

too-big-to-fail problem Problem in which regulators are reluctant to close down large financial institutions because doing so might precipitate a financial crisis.

trade balance on goods and services (trade balance) The difference between merchandise exports and imports.

transaction costs The time and money spent trying to exchange financial assets, goods, or services.

transmission mechanisms of monetary policy The channels through which the money supply affects economic activity.

Troubled Asset Relief Plan (TARP) A provision of the Bush administration's Economic Recovery Act of 2008 that authorized the Treasury to spend US$700 billion on purchases of subprime mortgage assets from troubled financial institutions or on injections of capital into these institutions.

trustee The representative of a person or corporation.

underfunded Describing a pension plan in which the contributions and their earnings are insufficient to pay out the benefits when they come due.

underwrite To purchase securities from a corporation at a predetermined price and then resell them in the market.

underwriting Guaranteeing the price for a corporation's securities and then selling them to the public.

unemployment rate The percentage of the labour force not working.

unexploited profit opportunity A situation in which an investor can earn a higher-than-normal return.

unit of account Anything used to measure value in an economy.

unsecured debt Debt not guaranteed by collateral.

unsterilized foreign exchange intervention A foreign exchange intervention in which a central bank allows the purchase or sale of domestic currency to affect the monetary base.

value at risk (VaR) Calculations that measure the size of the loss on a trading portfolio that might happen 1% of the time over a short period—say, two weeks.

vault cash Currency that is physically held by banks and stored in vaults overnight.

velocity of money The rate of turnover of money; the average number of times per year that a dollar is spent in buying the total amount of final goods and services produced in the economy.

venture-capital firm A financial intermediary that pools the resources of its partners and uses the funds to help entrepreneurs start up new businesses.

venture-capital funds Pooled investment funds managing the assets of investors who seek private equity stakes in high-risk and high-return startups as well as in small- to medium-sized enterprises with strong growth potential.

virtual bank A bank that has no building but rather exists only in cyberspace.

wealth All resources owned by an individual, including all assets.

World Bank The International Bank for Reconstruction and Development, an international organization that provides long-term loans to assist developing countries in building dams, roads, and other physical capital that will contribute to their economic development.

World Trade Organization (WTO) An organization headquartered in Geneva, Switzerland, that monitors rules for the conduct of trade between countries (tariffs and quotas).

yield curve A plot of the interest rates on particular types of bonds with different terms to maturity.

yield to maturity The interest rate that equates the present value of payments received from a credit market instrument with its value today.

zero-coupon bond See *discount bond*.

zero-lower-bound problem A situation in which the central bank is unable to lower short-term interest rates further because they have hit a floor of zero.

GUIDE TO COMMONLY USED SYMBOLS

Symbol	Term	Symbol	Term
Δ	change in a variable	i^*_{or}	target overnight interest rate
π	inflation rate	I	planned investment spending
π^e	expected inflation	IS	IS curve
π^T	inflation target	m	money multiplier
A	Bank of Canada advances	M	money supply
AD	aggregate demand curve	M^d	demand for money
AS	aggregate supply curve	M^s	supply of money
B^d	demand for bonds	$M1+$	M1+ monetary aggregate
B^s	supply of bonds	$M2+$	M2+ monetary aggregate
BR	borrowed reserves	MB	monetary base (high-powered money)
c	currency ratio	MB_n	nonborrowed monetary base
C	yearly coupon payment	MP	monetary policy curve
C	currency	mpc	marginal propensity to consume
C	consumption expenditure	NBR	nonborrowed reserves
D	demand curve	NX	net exports
D	chequable deposits	P	price level
DL	discount loans	P_s	stock prices
DR	desired reserves	P_t	price of a security at time t
e	excess reserves ratio	ρ	price shock
E_t	exchange (spot) rate	r	real interest rate
E_{par}	par (fixed) exchange rate	r_d	desired reserve ratio
$(E^e_{t+1} - E_t)/E_t$	expected appreciation of domestic currency	R	reserves
		R	return
EM	equity multiplier	R^e	expected return
ER	excess reserves	R^D	expected return on domestic deposits
$\bar{f}$	for financial frictions	R^F	expected return on foreign deposits
G	government purchases	ROA	return on assets
i	interest rate (yield to maturity)	ROE	return on equity
i_b	bank rate	S	supply curve
i_d	discount rate	T	taxes
i^D	interest rate on domestic assets	V	velocity of money
i^F	interest rate on foreign assets	Y	aggregate output (national income)
i_{er}	interest rate paid on reserves	Y^{ad}	aggregate demand
i_{or}	overnight interest rate	Y^P	potential (natural rate level of) output

Credits

Non-Prominent Image Credit(s):

A1 SHUTTERSTOCK: r.classen/Shutterstock; r.classen/Shutterstock; r.classen/Shutterstock; r.classen/Shutterstock; r.classen/Shutterstock; r.classen/Shutterstock; r.classen/Shutterstock; r.classen/Shutterstock; r.classen/Shutterstock; r.classen/Shutterstock; r.classen/Shutterstock; r.classen/Shutterstock; r.classen/Shutterstock; r.classen/Shutterstock; r.classen/Shutterstock; r.classen/Shutterstock

p. 1 SHUTTERSTOCK: r.classen/Shutterstock; r.classen/Shutterstock; r.classen/Shutterstock; r.classen/Shutterstock

p. 2 SHUTTERSTOCK: r.classen/Shutterstock

p. 19 SHUTTERSTOCK: r.classen/Shutterstock

p. 22 SHUTTERSTOCK: r.classen/Shutterstock

p. 48 SHUTTERSTOCK: r.classen/Shutterstock

p. 65 SHUTTERSTOCK: r.classen/Shutterstock

p. 87 SHUTTERSTOCK: r.classen/Shutterstock

p. 118 SHUTTERSTOCK: r.classen/Shutterstock

p. 143 SHUTTERSTOCK: r.classen/Shutterstock

p. 165 SHUTTERSTOCK: r.classen/Shutterstock

p. 189 SHUTTERSTOCK: r.classen/Shutterstock

p. 208 SHUTTERSTOCK: r.classen/Shutterstock

p. 243 SHUTTERSTOCK: r.classen/Shutterstock

p. 273 SHUTTERSTOCK: r.classen/Shutterstock

p. 301 SHUTTERSTOCK: r.classen/Shutterstock

p. 331 SHUTTERSTOCK: r.classen/Shutterstock

p. 355 SHUTTERSTOCK: r.classen/Shutterstock

p. 381 SHUTTERSTOCK: r.classen/Shutterstock

p. 424 SHUTTERSTOCK: r.classen/Shutterstock

p. 454 SHUTTERSTOCK: r.classen/Shutterstock

p. 477 SHUTTERSTOCK: r.classen/Shutterstock

p. 505 SHUTTERSTOCK: r.classen/Shutterstock

p. 523 SHUTTERSTOCK: r.classen/Shutterstock

p. 542 SHUTTERSTOCK: r.classen/Shutterstock

p. 557 SHUTTERSTOCK: r.classen/Shutterstock

p. 592 SHUTTERSTOCK: r.classen/Shutterstock

p. 617 SHUTTERSTOCK: r.classen/Shutterstock

p. 636 SHUTTERSTOCK: r.classen/Shutterstock

p. S1-1 SHUTTERSTOCK: r.classen/Shutterstock

p. S2-1 SHUTTERSTOCK: r.classen/Shutterstock

p. S3-1 SHUTTERSTOCK: r.classen/Shutterstock

p. S3-20 SHUTTERSTOCK: r.classen/Shutterstock

Prominent Image Credit(s):

PopTika/Shutterstock

Index

Note: Page numbers followed by *b*, *f*, or *t* represent boxes, figures, or tables, respectively.

A

Activists, 598, 600*b*
Actual expenditure, 523
Adaptive expectations, 149
Adjusted forward-rate forecast, 138
Advances, 275
Adverse selection, 170–176, 289
  collateral, 175
  definition of, 170
  financial intermediaries, 36–38
  financial intermediation, 173–175
  government regulation to increase information, 172–173
  government safety net and, 192
  net worth, 175
  private production, 172
  sale of information, 172
  summary, 175–176
After-tax real interest rate, 82*n*
Agency theory, 170
Aggregate demand, 523
  autonomous consumption expenditure, 564
  autonomous investment, 566
  autonomous net exports, 564
  business cycles, 557–560
  components of, 524–529, 561–562
  conclusions from, 584–588
  consumption expenditure, 524–525
  equilibrium in, 573–578
  financial frictions, 566
  government purchases, 527, 564
  inflation, 560, 560*f*
  IS curve, 523–529
  net exports, 528–529
  planned expenditure and, 523–524
  planned investment spending, 525–527
  taxes, 527–528, 564
Aggregate demand–aggregate supply (AD/AS) model, 523, 574, 575, 585–587
  Covid-19 recession in Canada, 587–588
Aggregate demand curve
  autonomous monetary policy tightening, 553*f*
  definition of, 548, 562
  deriving, 548–550, 550*b*, 562
  with effective lower bound, 605–607, 606*f*
  shift in, 562–566, 563*f*, 564*f*, 565*t*
  shifts in *IS* curve, 551, 552*f*
  shifts in *MP* curve, 551, 554
Aggregate demand shocks, 578–581
  credibility and, 624–626, 625*f*
  monetary policy, 593–594, 594*f*, 595*f*
Aggregate income, 19
Aggregate output, 19, 557, 561
  index, 574–576, 576*f*
  money and, 7
Aggregate price level, 19
  money and, 8*f*
Aggregate supply, 566–569
  business cycles, 557–560
  conclusions from, 584–588
  equilibrium in, 573–578
  inflation, 560, 560*f*
Aggregate supply curve
  definition of, 566
  expected inflation, 567–568
  inflation shocks, 568
  long-run, 566–567, 567*f*
  output gap, 568, 571–572, 572*f*
  shifts in, 569–573, 570*f*, 571*f*, 573*t*
  short-run, 567–569, 567*f*
  supply shocks, 568
Aggregate supply shocks, 581–584
  credibility and, 626–627, 626*f*
Agricultural Bank of China, 236*t*
Akerlof, George, 170, 171*n*
Alberta Municipal Financing Corporation, 31
Alberta Treasury Branches (ATB) Financial, 233
American options, 311
American Recovery and Reinvestment Act of 2009, 258
Anchor country, 494
Anchor currency, 482
Animal spirits, 527, 551
Annual percentage rate (APR), 199
Anti-inflation policy, 629–631, 629*f*, 630*b*
Apple, 22
Appreciation, 455
Arbitrage, 153, 305
Argentina's currency board, 499*b*
Asset-backed commercial paper (ABCP), 260*b*–261*b*
Asset demand, determinants of, 87–89
Asset–liability management committee, 284
Asset management, 280, 283–284
Asset market approach, 93
Asset-price boom, 246
Asset-price bubbles, 246
  bubbles driven solely by irrational exuberance, 440
  Central banks, 439–442
  credit-driven bubbles, 439
  definition of, 439
  macroprudential policies, 441–442
  monetary policy, 442

Asset-price bust, 246
Assets, 222–223
  bank balance sheet, 275–277
  Bank of Canada's balance sheet, 357–358
  decline of traditional banking, 222–223
  definition of, 3
Asset transformation, 36, 277
Asymmetric information, 170, 182*t*
  adverse selection, 170–176
  financial intermediaries, 36–38
  financial regulation, 189–193
  global financial crisis, 251
  lemons problem, 170–176
  moral hazard, 170, 176–181
  principal–agent problem, 176–177
Automated Clearing Settlement System (ACSS), 383
Automated teller machines (ATM), 216, 228
Autonomous consumption, 536
Autonomous consumption expenditure, 525
  aggregate demand, 564
Autonomous easing of monetary policy, 545
Autonomous investment, 526
  aggregate demand, 566
  spending, 536
Autonomous meaning, 563*b*
Autonomous monetary policy, 563
Autonomous net exports, 528, 75, 36
  aggregate demand, 564
Autonomous spending, *IS* curve, 536–537
Autonomous tightening of monetary policy, 545

B

Baa, 122–123
Balance of payments, 481–482
  current account, 481
  financial account, 481–482
Balance sheet channel, 641–642
Balance sheets, 273–277
  Bank of Canada, 356–358
  definition of, 273
  financial institutions, 254
  Great Depression and, 644*b*
Bank Act, 1881–1913, 212–213
Bank Act reform, 237–240
  bank holding companies, 237–238
  Canadian banking industry, 240
  Canadian Payments Act (CP Act), 239
  definition of, 237
  merger review policy, 239–240
  National Financial Services OmbudService (NFSO), 240
  ownership rules, 238–239
  permitted investments, 238
Bank-as-parent model, 237
Bank audit, 213
Bank balance sheet, 273–277
  assets, 275–277
  liabilities, 273–275
Bank capital, 275
Bank capital requirements
  capital adequacy management, 287
Bank commercial loans, 31
Bankers' Acceptance Purchase Facility, 406, 407*b*
Banker's risk, 276
Bank failure, 285
  definition of, 189
Banking industry
  Bank Act reform, 237–240
  Canadian banking industry, 208–213
  Canadian chartered banking industry, 224–225
  convergence, 228–229
  cooperatives, 232–233
  financial consolidation, 229
  financial services industries, 229–230
  four-pillar approach, 228–230
  global financial crisis, 230*b*
  government savings institutions, 233
  international banking, 233–236
  loan companies, 231–232
  regulation and structure, 231–233
  restrictions on branching, 225–227
  shadow banking system, 213–223
  trust companies, 231
  United States comparison, 225–227
Bank lending channel, 641
Bank loans, 165
Bank management, 280–288
  asset management, 280, 283–284
  capital adequacy management, 280, 285–288
  credit risk, 280
  deposit outflows, 280
  interest-rate risk, 280
  liability management, 280, 284
  liquidity management, 280–283
Bank of America Corp., 226*t*, 236*t*, 256
Bank of Canada, 10, 282, 362, 369
  accountability and trans parency, 340
  balance sheet, 356–358
  bank's research staff role, 336*b*
  Board of Directors, 332
  case against independence, 342–343
  case for independence, 341–342
  central bank independence, 343
  changing face of, 338–340
  conventional monetary policy, 402, 402*f*
  credibility and, 631*b*
  currency, 334
  disinflation, 580–581, 581*f*
  economics and politics, 342*b*
  financial system, 334–335

Bank of Canada (*Continued*)
functions of, 333–336
funds management, 334
Governing Council, 333
independence of, 336–340
macroeconomic performance, 343
monetary aggregates, 56–59
monetary base, 362
monetary conditions, 391
monetary policy and, 335, 390–395, 542–543
open market operations, 358–359
origins of, 331–332
policy rate, 383–384
political environment and, 333*b*
standing facilities, 386
Taylor rule, 448*b*
Bank of Canada Act in 1934, 332
Bank of Canada lending, 402, 402*f*
advantages of, 404
disadvantages of, 404
Bank of China, 236*t*
Bank of England, 351
Bank of Japan, 351–352
Bank panics
definition of, 190
financial crisis, 248–249
government safety net, 189–191
Bank rate, 358
Banks, 14, 15
asset transformation, 277
definition of, 6
dual banking system in United States, 210*b*
money supply, 355
national banks, 210*b*
responses, 223
state banks, 210*b*
T-account, 277–278
U.S. banks, 226*t*
Barter economy, 49
Basel Accord
after global financial crisis, 196*b*
definition of, 194
Basel Committee on Bank Supervision, 194
Base money, 334–335
Basic gap analysis, 293
Bearer deposit notes, 28
Behavioural finance, 159–160
Bell Canada bond, 88
Big Macs, 459–461, 460*t*
Bitcoin, 54–55
BNP Paribas, 236*t*
Board of Directors, 332
Board of Governors, 344
Board of Governors of Federal Reserve System, 344
Bolivian hyperinflation, 630*b*
Bombardier, 29
Bond markets, 3
lemons problem in, 171
supply and demand in, 90–93, 91*f*
Bonds, 23, 165
corporate, 30–31
coupon bond, 68, 72–75, 73*t*
definition of, 3
discount bond, 75–76
Government of Canada Bonds, 31
junk bonds, 122, 217
over-the-counter (OTC) market, 26
ratings, 121*t*
shifts in demand for, 93–97, 94*f*
shifts in supply of, 97–103, 97*t*, 98*f*
Booms, 557
Borrowed reserves (BR), 358, 362, 369, 388
Borrowings, 275
Branches, 209
Branching, restrictions on, 226–227
Bretton Woods system, 483–484
Brexit (Britian Exit), 454, 473–474
British pounds, 473–474, 486*f*
British-style universal banking system, 230
Brokers, 26
Bubbles, 158
Budget deficit, 11
inflation and, 511–514
Budget surplus, 11, 12*f*
Bull market, 4
Business cycles
inflation and, 557–560
money and, 7–8
political, 621*b*

C

Caisses populaires, 6, 30, 203, 208
convergence, 229
credit unions and, 232–233
money supply and, 57
Call, 31
Call option, 312, 312*f*
Call premium, 312
CAMELS rating, 197
Canada
AD/AS analysis of Covid-19 recession in, 587–588
balance sheet of all banks in, 274*t*
global financial crisis in, 260–263
inflation rates and inflation targets for, 392*f*
inflation targeting in, 430–432, 431*f*
subprime mortgages in, 261–262
Canada Deposit Insurance Corporation (CDIC), 44, 203–205
bail-in power, 205
differential premiums, 203–204, 204*t*
opting out, 204
resolution authority, 205
Canada Mortgage and Housing Corporation (CMHC), 30, 191
Canada Mortgage Bond Purchase Program, 406, 407*b*
Canadas, 31
Canadian acquisition of financial assets, 481

Canadian banking industry
Bank Act, 1881–1913, 212–213
Dominion Notes Act, 211–212
Finance Act of 1914, 213
First Bank Act, 212
free banking experiment, 210–211
historical development, 208–213, 209*f*
implications for, 240
Provincial Notes Act, 211
Canadian banking overseas, 235
Canadian banks, 233, 262–263
Canadian chartered banking industry, 224–225, 224*t*
competition and technology, 225
Schedule I banks, 224–225
Schedule II banks, 224–225
Schedule III banks, 224–225
Canadian Payments Act (CP Act), 239
Capital, 24
Capital adequacy management, 280, 285–288
bank capital requirements, 287
equity multiplier (EM), 286
prevent bank failure, 285
return on assets (ROA), 286
return on equity (ROE), 286
strategies for, 287
trade-off between safety, 286–287
Capital controls, 234, 489
inflows, 491–492
international financial system, 491–492
outflows, 491
Capital market
bank commercial loans, 31
consumer loans, 31
corporate bonds, 30–31
definition of, 27
government agency securities, 31
Government of Canada Bonds, 31
instruments, 29–31, 30*t*
interest rates, 30*b*
mortgage-backed securities, 30
mortgages, 30
municipal government bonds, 31
provincial bonds, 31
stocks, 29
Capital mobility, 485
Capital One Financial Corp., 226*t*
Cash flows
channel, 642–643
definition of, 65, 144
Cash items, 276
Cease and desist orders, 197
Central banks, 10, 209, 332*b*
asset-price bubbles, 439–442
Bank of England, 351
Bank of Japan, 351–352
credibility and, 631–632
European Central Bank (ECB), 349–351
Federal Reserve System, 344–349
independence, 343
money supply, 355
theory of bureaucratic behaviour, 343–344
trend toward greater independence, 352
Certificates of deposit (CDs), 28, 275
Ceteris paribus, 90
Channel, 384
monetary control in, 400–401
Chartered banks, 39, 195, 208*n*
Chartering financial institutions, 195
Cheques, 52–53
China, international reserves in, 489
China Construction Bank Corp., 236*t*
Citigroup Inc., 226*t*
Collateral
adverse selection, 175
credit risk, 291
definition of, 168
moral hazard, 179–180
Collateralized debt obligations (CDOs), 250, 251*b*
Commercial paper
definition of, 29
information technology, 217–218
money market, 29
Commercial Paper Purchase Program, 406, 407*b,* 408
Commodity money, 52
Common stock, 3
calculate price of, 143–146
definition of, 143
generalized dividend valuation model, 145
Gordon growth model, 145–147
one-period valuation model, 144–145
Compensating balances, 291
Consol, 74
Constant-money-growth-rate rule, 620
Constitution Act, 211
Constrained discretion, 622–623
Consumer durable expenditure, 637
Consumer expenditure, 640
Consumer loans, 31
Consumer price index (CPI), 20, 390, 561
Consumer Protection Act of 2010, 264–266
Consumer reports, 174
Consumption, 640
Consumption expenditure, 524
Consumption function, 524–525
Contagion effect, 190
Contingent Term Repo Facility, 407*b,* 408
Contractual savings institutions, 40–41
Conventional monetary policy, 395–405
Bank of Canada lending, 402, 402*f*
definition of, 395
open market (buyback) operations, 395–397
settlement balances management, 398–400

Convertible bonds, 30–31
Cooperatives, 232–233
Coronavirus crisis, monetary policy curve and, 547–548
Coronavirus pandemic, 122–123
Coronavirus stock market crash, 148
Corporate Bond Purchase Program, 407*b,* 408
Corporate bonds
  capital market instruments, 30–31
  default risk, 119, 120*f*
Correspondent banking, 276
Corridor
  definition of, 384
  monetary control in, 400–401
Costly state verification, 177
Cost-push inflation, 602, 603*f*
  definition of, 602, 603*f*
  vs. demand-pull inflation, 605
Cost-push shocks, 568
Counterparty, 303
Coupon bond
  consol/perpetuity, 74
  definition of, 68
  interest-rate risk, 80*n*
  yield to maturity, 72–75, 73*t*
Coupon rate, 68
Covid-19 Pandemic, 407*b*
Covid-19 recession in Canada, 587–588
Coyne, James, 337
Crawling target/peg, 494
Credibility, 623–631
  aggregate demand shocks and, 624–626, 625*f*
  aggregate supply shocks and, 626–627, 626*f*
  anti-inflation policy and, 629–631, 629*f,* 630*b*
  appoint conservative Central bankers, 632
  Bank of Canada and, 631*b*
  benefits of, 623–624
  Central bank, 631–632
  definition of, 623
  inflation and unemployment, 628*f*
  nominal GDP targeting, 632
  oil price shocks, 627
Credit Agricole Group, 236*t*
Credit boom, 244–246
Credit bust, 244–246
Credit channels, 644–645
Credit default swap, 251, 324
Credit derivatives, 323–325
  credit-linked notes, 324
  credit options, 323
  credit swaps, 323–324
  definition of, 323
Credit-driven bubbles, 439
Credit easing
  definition of, 411
  quantitative easing *vs.,* 410–411
Credit event, 324
Credit-linked notes, 324
Credit market instruments, 68–69
Credit options, 323
Credit-rating agencies, 251
  conflicts of interest, 122*b*
  definition of, 121
Credit rationing, 291–292
Credit risk
  collateral, 291
  compensating balances, 291
  credit rationing, 291–292
  definition of, 280
  loan commitments, 291
  long-term customer relationships, 290
  management, 280, 288–292
  monitoring, 290
  restrictive covenants, 290
  screening, 289
  specialization in lending, 289–290
Credit spread, 249, 527
Credit swaps, 323–324
Credit unions, 232–233
Credit Unions and *Caisses Populaires* (CUCPs), 39
Credit view
  balance sheet channel, 641–642
  bank lending channel, 641
  cash flow channel, 642–643
  definition of, 638
  household liquidity effects, 643–644
  unanticipated price level channel, 643
Cryptocurrencies, 54–55
Currency, 334, 356
  definition of, 28, 48
  M1+, 57
  shifts from deposits into, 359–360
Currency board, 498, 499*b*
Currency swaps, 320
Currency union, 489–490
Current account, 481
Current account balance, 481
Current yield, 75

D

Data lag, 599
Dealers, 26
Debit cards, 53
Debt deflation, 249–250
Debt markets, 3, 25
  maturity, 25
Default, 119
  definition of, 28
Default-free bonds, 119
Deleveraging, 244
Demand curve, 387–388
  definition of, 90
  for domestic assets, 464–465
Demand deposits, 274
Demand for assets, determinants of, 87–89
Demand for money
  definition of, 507
  interest rates and, 518–519
  stability of, 519
Demand-pull inflation, 603–604, 604*f*
  definition of, 603
  vs. cost-push inflation, 605
Demand shocks, 562
Deposit facility, 420
Deposit insurance, 44, 189–191, 191*b*
Depositors, 355

Depository institutions, 38–40
Deposit outflows, 280
Deposit rate ceilings, 220
Deposit receipts (DRs), 28
Depreciation, 455
Derivatives, 266, 268
Desired reserve ratio, 276, 357, 369–370
Desired reserves, 276, 357
Devaluation, 485
Direct clearers, 383
Direct finance, 23, 23*f*
Dirty float, 482
Discount bond, 69
  yield to maturity, 75–76
Discount lending, 417–418, 417*f*
Discount rate, 282, 414
Discount window, 414
Discretion, 619–622
Disintermediation, 220, 222
Disposable income, 524
Diversification, 36
Dividends, 25, 144
Dodd-Frank wall street reform, 264–266
Dollarization, 499–500
Domestic assets
  demand curve for, 464–465, 469*t*
  shifts in demand for, 465–468
  supply curve for, 463
Domestic interest rate, 465*f,* 466
Dominated assets, 517
Dominion Notes Act of 1870, 211–212
Drawdowns, 398
Dual banking system in United States, 210*b*
Dual mandates
  definition of, 429
  vs. hierarchical mandates, 428–429
Duration analysis, 293–294

E

E-cash, 53–54
Econometric policy evaluation, 618
Economic Stimulus Act of 2008, 258
Economies of scale, 35, 169
Economies of scope, 38, 229
Effectiveness lag, 599
Efficient market hypothesis, 143, 151
  behavioural finance, 159–160
  definition of, 152
  development of, 152*n*
  for financial markets, 158–159
  foreign exchange market, 155*b*
  investment advisor reports, 155–156, 157*b*
  market efficiency and, 157–158
  random-walk behaviour of stock prices, 154–155, 155*b*
  rationale of, 153–154
  short sales, 160
E-finance, 6
Electoral business cycles, 342*b*
Electronic money, 53–55
Electronic payment, 53
Emergency Economic Stabilization Act, 257
Emergency lending assistance to troubled banks, 405*b*
Emerging market countries, 491
E-money, 53–55, 54*b*
Enron Corporation, 173*b*
Equilibrium interest rate
  changes in, 93–103
  liquidity preference framework, 106–109
  shifts in demand for bonds, 93–97, 94*f*
  shifts in supply for bonds, 97–103, 97*t,* 98*f*
Equilibrium price, 92
Equities, 25
Equity capital, 175
Equity markets, 25
Equity multiplier (EM), 286
Equity tranche, 251*b*
Eurobonds, 33
Eurocurrencies, 33, 234–235
Eurodollars, 33
European Central Bank (ECB), 349–351
  excess reserves, interest on, 420
  Federal Reserve System, 349
  Governing Council, 349–350
  independence of, 350–351
  lending to banks, 420
  monetary policy, 420–421, 432*b*
  open market operations, 420
  reserve requirements, 421
European Monetary System (EMS), 484
European options, 311
European sovereign debt crisis, 256*b*
European System of Central Banks (ESCB), 349
Euro survive, 490*b*
Eurosystem, 349, 350
Ex ante real interest rate, 81
Excess demand, 92
Excess reserves, 357
  interest on, 419, 419*f,* 420
Excess supply, 92
Exchange Fund Account, 334, 399–400
Exchange rate, 455–456, 528
Exchange rate mechanism (ERM), 484
Exchange-rate peg, 494
Exchange-rate targeting, 494–500, 631
  advantages of, 494–495
  currency board, 498, 499*b*
  disadvantages of, 495–497
  dollarization, 499–500
  for emerging market countries, 498
  for industrialized countries, 497
Exchanges, 26
Exercise price, 311
Exogenous, 525
Expansions, 557
Expectations theory, 127–130
  definition of, 127
  investment strategies, 128
  liquidity premium theory and, 133*f*

Expectations theory (*Continued*)
term structure of interest rates, 127–130
volatility and, 130*n*
Expected future exchange rate, 467–468, 468*f*
Expected inflation
effect, 110, 111
equilibrium interest rate, 98, 100*f*
Expected profitability of investment opportunities, 98
Expected returns
asset demands and, 88, 89*t*
bond demand and, 95*t*, 96
definition of, 88
Ex post real interest rate, 81

F

Face value, 68
Fallen angels, 217
Favourable supply shock, 582
Federal Deposit Insurance Corporation (FDIC), 190, 405
Federal Deposit Insurance Corporation (FDIC)-insured banks, 226*t*
Federal funds, 414
Federal funds rate, 346, 414
Federal Open Market Committee (FOMC), 344, 346
Federal Reserve, 265
discount lending, 417–418, 417*f*
discount rate, 414
discount window, 414
excess reserves, interest on, 419, 419*f*
federal funds rate, 414
inflation targeting, 438*b*
market for reserves in United States, 415
monetary policy, 414–419
open market operations, 416–417, 416*f*
primary credit, 414
reserve requirements, 418, 418*f*
seasonal credit, 414
secondary credit, 414
Federal Reserve Banks, 344, 347*b*
Federal Reserve System, 344–349, 346*f*
Board of Governors, 344
European Central Bank (ECB), 349
Federal Open Market Committee, 346
Federal Reserve bank, 344–345
independence of, 346–349, 348*b*
policy tools in, 345*f*
Fiat money, 52
Finance Act of 1914, 213
Finance companies, 41
Finance paper, 29
Financial account, 481–482
Financial account balance, 481
Financial banking, 5–7
Financial crisis, 7, 243
in advanced economies, 245
banking crisis, 246–247
Consumer Financial Protection Bureau, 267
debt deflation, 247–248
definition of, 244
derivatives, 268
dynamics of, 244–250
global financial crisis, 250–259
Government-Sponsored Enterprises (GSEs), 268
initial phase, 244–246
resolution authority, 267–268
too-big-to-fail problem, 266–267
Volcker rule, 268
Financial derivatives, 215, 251
definition of, 301
global financial crisis, 324–25
Financial development
China, 184–185
and economic growth, 182–184
Financial engineering, 214
Financial frictions, 527
aggregate demand, 566
definition of, 244
IS curve, 538
Financial futures contract, 304–311
definition of, 304
hedging foreign exchange risk with, 310–311
vs. options, 316*f*
Financial futures markets, 307–309
globalization of, 307–308
organization of trading, 307, 307*t*
success of, 308–309
Financial futures options, 311
Financial innovation, 6–7
decline of traditional banking, 221–223
definition of, 244
mortgage markets, 250
shadow banking system, 213–223
Financial institutions, 5–7, 14
balance sheets, 254
loans to, 357–358, 360–361
too-big-to-fail problem, 266
Financial instruments, 3, 23
Financial intermediaries, 6, 22, 40*t*
asymmetric information, 36–38
conflicts of interest, 38
contractual savings institutions, 40–41
depository institutions, 38–40
economies of scope, 38
investment intermediaries, 41–42
primary assets and liabilities of, 39*t*
risk sharing, 36
soundness of, 43–44
transaction costs, 34–35
Financial intermediation, 34, 35*f*, 173–175
Financial legislation in Canada, 201*t*–202*t*
Financial liberalization, definition of, 244
Financial markets, 14, 22
definition of, 2
financial banking, 5–7
financial institutions, 5–7
function of, 22–24, 23*f*
instruments, 27–31
international finance, 12–14, 13*f*
stability of, 427
structure of, 25–27

Financial panic, 43
Financial regulation
  assessment of risk management, 197–198
  asymmetric information, 189–193
  Canada Deposit Insurance Corporation (CDIC), 203–205
  capital requirements, 194–195
  consumer protection, 199
  Consumer Protection Act of 2010, 264–266
  disclosure requirements, 198–199
  Dodd-Frank wall street reform, 264–266
  financial legislation in Canada, 201*t*–202*t*
  financial supervision, 195–197
  global financial crisis, 263–266
  government safety net, 189–193
  international, 200*b*
  macroprudential *vs.* microprudential supervision, 263–264
  prompt corrective action, 195
  restrictions on asset holdings, 193–194
  restrictions on competition, 199
  types of, 193–202
Financial repression, 182
Financial services industries, 229–230
Financial structure. *See* Financial system
Financial supervision, 195–197
Financial system. *See also* Financial markets
  adverse selection, 170–176
  Bank of Canada, 334–335
  basic facts of, 165–168, 166*f*
  Eurobonds, 33
  Eurocurrencies, 33
  financial intermediation, 34, 35*f*
  financial regulation abroad, 44–45
  foreign bonds, 33
  indirect finance, 33–38
  international bond market, 32–33
  internationalization of, 31–33
  lemons problem, 170–176
  moral hazard, 170
  regulation of, 42–45
  regulatory agencies of Canadian financial system, 43*t*
  structure of, 6
  transaction costs, 168–169
  world stock markets, 33
Fire sales, 247
First Bank Act, 212
Fiscal policy, 11–12, 258
Fisher effect, 99–100
Fisher equation, 81
Fixed deposit. *See* Certificate of deposit (CD)
Fixed exchange rate regime, 482, 484–485, 484*f*
Fixed investment, 525
Fixed-payment loan
  definition of, 68
  yield to maturity, 70–71
Fixed-term deposits, 274–275
Floating exchange rate regime, 482
Floor system of monetary control, 409*b*
Ford Motor Company, 23, 32*b*, 525
Foreign banks in Canada, 235–236
Foreign bonds, 33
Foreign-exchange crisis, 486–487
Foreign exchange interventions
  definition of, 477
  and money supply, 477–479
  sterilized intervention, 479, 480
  unsterilized intervention, 479–480, 480*f*
Foreign exchange market, 13–14, 454
  Brexit, 473–474
  British pound, 473–474, 486*f*
  demand curve for domestic assets, 464–465, 469*t*
  equilibrium in, 464*f*, 465
  exchange rates, 455–456
  factors affect, 461–463, 462*t*
  global financial crisis, 472–473
  interest rates, 470–472
  intervention in, 477–480
  monetary policy, 493–494
  purchasing power parity (PPP), 457–461
  stability in, 428
  supply and demand analysis, 463–465
  supply curve for domestic assets, 463
  trades, 456
  U.S. dollar, 472–473
Foreign exchange rates, 13, 455
Foreign interest rate, 466–467, 467*f*
Foreign stock market indexes, 34*b*
Forward contracts
  definition of, 302
  hedging foreign exchange risk with, 310
  pros and cons of, 303
Forward exchange rate, 455
Forward transactions, 455
Four-pillar approach, 228–230
Free banking, 210–211
Free-rider problem, 172
Free-standing investment banks, 230*b*
Frictional unemployment, 426
Friedman, Milton, 9
Fully amortized loan, 68. *See also* Fixed-payment loan
Fundamental economic values, 246
Funds management, 334
Futures contracts, 215
Futures options, 311, 314–317

G

Gap analysis, 293–294
General Agreement on Tariffs and Trade (GATT), 483
General Electric, 151

Generalized dividend valuation model, 145
General Motors, 29, 32*b*
Glass-Steagall, 230*b*
Global financial crisis, 250–259
  agency problems in mortgage markets, 250–252
  asymmetric information, 251
  Canada, 260–263
  coronavirus pandemic, 259
  credit crunch, 288
  credit-rating agencies, 251
  failure of high-profile firms, 255–257
  financial derivatives, 324–25
  financial innovation in mortgage markets, 250
  financial institutions' balance sheets, 254
  financial regulation, 263–266
  foreign exchange market, 472–473
  global financial markets, 255
  government bailouts, 258*b*
  government intervention and recovery, 258–259
  monetary policy, 436–438
  monetary policy (MP) curve, 547–548
  residential housing prices, 252–254, 252*f,* 253*b*
  shadow banking system, 254–255, 255*f*
Global financial markets, 255
Goal independence, 336, 339
Goldman Sachs Group, 226*t*
Gold standard, 211, 483
Goods market equilibrium, 529–530
  IS curve, 530
  solving for, 529–530
Gordon growth model
  definition of, 145
  stock prices, 147
Governing Council, 339
  Bank of Canada, 333
  European Central Bank (ECB), 349–350
Government budget constraint, 511–512
Government budget deficits, 98–99
Government of Canada Bond Purchase Program, 406, 407*b*
Government of Canada Bonds, 31
  interest rates on, 127*f*
  yield curve for, 136*f*
Government of Canada Treasury Bills, 28
Government purchases, 524
  aggregate demand, 527, 564
  IS curve, 532–533, 533*f*
Government retirement funds, 41
Government safety net, 189–193
  adverse selection and, 192
  bank panics, 189–191
  deposit insurance, 189–191, 191*b*
  drawbacks of, 191–193
  financial consolidation and, 193
  moral hazard and, 192
  other forms of, 191
  too-big-to-fail problem, 192–193
Government savings institutions, 233
Government-Sponsored Enterprises (GSEs), 268
Great Depression, 7, 193, 229
  balance sheets and, 644*b*
  Bank of Canada, 331
  bank panics, 248–249
  debt deflation, 249–250
  international dimensions, 250
  stock market crash, 248
  stock prices, 249
  in United States, 248–250
Great Recession, 7, 257, 585–587, 645
Gross domestic product (GDP), 11, 12*f,* 19
  deflator, 20
  and employment growth, 558*f*
  nominal GDP targeting, 632
  Taylor rule, 446
Guaranteed Investment Certificates (GICs), 28
Guaranteed Investment Receipts (GIRs), 28
Guaranteed Trust Certificates (GTC), 28

H

Haircuts, 254
Happy Feet Corporation (HFC), 154, 156
Hedge, 215, 301
Hedge funds, 42
Hedging
  financial futures, 305–306
  with futures options, 317–318
  with interest-rate forward contracts, 302–303
  interest-rate swaps, 321–322
  long position, 301
  short position, 301
Hedging foreign exchange risk, 309–311
  with forward contracts, 310
  with futures contracts, 310–311
Helicopter money, 413*b*
Hierarchical mandates
  definition of, 428
  vs. dual mandates, 428–429
High-powered money, 356, 371
Holding companies, 227
  Bank Act reform, 237–238
Home banking, 216
Honda, 32*b*
Household liquidity effects, 643–644
HSBC Holdings PLC, 236*t*
Hyperinflation, 51, 424, 513–514
  Bolivian, 630*b*
  definition of, 513
  Zimbabwean, 514

I

Implementation lag, 599
Impossible trinity, 488
Incentive-compatible, 180
Income, 19, 49
  changes in, 108
  income effect, 106, 110
Income effect, 106, 110
Income tax considerations, 124
Indebtedness, 235
Indirect finance, 23*f*
Industrial and Commercial Bank of China, 236*t*

Inflation, 561
budget deficits and, 511–514
business cycles and, 560, 560*f*
high employment targets and, 602–605
money and, 8–10, 10*f*
Inflation gap, 592
Inflation hedges, 517
Inflation rate, 9, 21, 546*f*
Inflation-rate targeting, price-level targeting *vs.,* 390
Inflation report, 433
Inflation shock, 570–571
Inflation targeting, 429–435, 592, 631
advantages of, 433–434
in Canada, 430–432, 431*f*
consistency with democratic principles, 434
definition of, 429
disadvantages of, 434–435
Federal Reserve, 438*b*
flexibility of, 438
implications for, 437–438
increased accountability, 433–434
increased transparency, 433
level of, 437–438
in New Zealand, 430, 431*f*
performance of, 434
reduction of time-inconsistency problem, 433
in United Kingdom, 432
Information technology
bank credit, 215–216
commercial paper, 217–218
debit cards, 215–216
definition of, 215
electronic banking, 216–217
home banking, 216
junk bonds, 217
Instrument independence, 336
Interest parity condition, 465*n*
Interest-rate forward contracts
definition of, 302
hedging with, 302–303
Interest-rate risk, 214
coupon bond, 80*n*
definition of, 79, 280
duration of, 79*n*
gap and duration analysis, 293–294
management, 292–294
strategies for, 294
Interest rates, 65
asset demand, determinants of, 87–89
business cycle expansion, 101–102, 101*f,* 102*f*
capital market, 30*b*
credit market instruments, 68–69
definition of, 3
demand for money and, 518–519
equilibrium interest rate, 93–103
foreign exchange market, 470–472
on Government of Canada Bonds with maturities, 127*f*
measurement of, 65–76
money and, 10, 11*f,* 109–113
on municipal and treasury bonds, 125*f*
negative, 76*b*
nominal, 81–83, 83*f*
present value, 65–68
real, 81–83, 83*f*
and returns, 76–81, 79*t*
risk structure of (*see* Risk structure of interest rates)
on selected bonds, 4*f*
supply and demand, in bond market, 90–93, 91*f*
term structure of (*see* Term structure of interest rates)
Interest-rate stability, 427
Interest-rate swaps, 320–322
advantages of, 321–322
disadvantages of, 322
financial intermediaries in, 322
hedging with, 321–322
payments, 320*f*
Intermediate target, 443
Intermediate term, 25
International banking
Canadian banking overseas, 235
Eurocurrencies market, 234–235
foreign banks in Canada, 235–236
ironic birth of Eurodollar market, 234*b*
world largest banks, 236*t*
International bond market, 32–33
International finance, 12–14, 13*f*
International financial regulation, 200*b*
International financial system
balance of payments, 481–482
Bretton Woods system, 483–484
capital controls, 491–492
exchange rate regimes in, 482–491
exchange-rate targeting, 494–500
fixed exchange rate regime, 484–485, 484*f*
foreign exchange interventions, 477–480
gold standard, 483
International Monetary Fund, 492–493
managed float, 490–491
monetary policy, 493–494
monetary unions, 489–490
policy trilemma, 488–489, 488*f*
speculative attack, 486
International Monetary Fund (IMF), 483, 492–493
International reserves, 477
Inventory investment, 525–526
Inverted yield curve, 125
Investment, 525*b*
Investment bank, 26, 42
Investment intermediaries, 41–42
IS curve, 530, 537*t*
aggregate demand, 523–529
autonomous spending, 536–537
economy heads, 532

IS curve (*Continued*)
financial frictions, 538
goods market equilibrium, 529–530
government purchases, 532–533, 533*f*
intuition, 530
numerical example, 530–531, 531*f*
planned investment spending, 524–527
shifts in, 551, 552*f*
taxes, 534–535

J

Japan
monetary policy lessons from, 647–648
J.P. Morgan Chase & Co., 226*t*, 236*t*
Junk bonds, 122, 217

K

Keynes, John Maynard, 103
Keynesians, 598
Keynesian theories of money demand, 514–516
liquidity preference theory, 514
precautionary motive, 515
real money balances, 515
speculative motive, 515
transactions motive, 515

L

Lack of liquidity, 303
Large-scale asset purchases, 408–409, 408*f*
Large Value Transfer System (LVTS), 357, 382, 385*b*
Legislative lag, 599
Lemons problem, 170–176
Lender of last resort, 213, 403–404, 404*f*
Leverage cycle, 264
Leverage ratio, 194
Liabilities, 23, 222
bank balance sheet, 273–275
Bank of Canada's balance sheet, 356–357
decline of traditional banking, 222
management, 280, 284
Life insurance companies, 40–41
Liquid, 26
Liquidity, 51, 303, 517
asset demands and, 89, 89*t*
bond demand and, 95*t*, 96–97
definition of, 88
risk structure of interest rates, 123
Liquidity effect, 110, 111
Liquidity management, 280–283
Liquidity preference framework, 103–106
equilibrium interest rates in, 106–109
income, changes in, 108
money supply, changes in, 108–109, 109*f*
price level, changes in, 108
Liquidity preference theory, 514
Liquidity premium theory, 131–135
definition of, 131
expectations theory and, 133*f*
yield curve, 134*f*
Liquidity provision, 406–408, 609
Liquidity services, 35, 169
Liquidity trap, 518
Loan commitments, 291
Loan companies, 39, 231–232
Loan originator, 218
Loan sales, 295
Longer-term refinancing operations, 420
Long in a call, 312
Long position, 301, 304
Long term, definition of, 25
Long-term interest rate, 393
Lucas critique, 617–618

M

Macroeconometric models, 617
Macroeconomic equilibrium, 573
Macro hedge, 306
Macroprudential policies, 441–442
Macroprudential regulation, 442
Macroprudential supervision, 263–264
Main refinancing operations, 420
Managed float, 490–491
Managed float regime, 482
Management of expectations, 412
Marginal lending facility, 420
Marginal lending rate, 420
Marginal propensity to consume, 525
Margin requirement, 309
Marked to market, 309
Market-clearing price, 92
Market equilibrium, 87, 92–93
Market for money
definition of, 103*n*
equilibrium in, 105*f*
Market fundamentals, 158
Maturity, 25
Maturity bucket approach, 293
Maturity date, 66
McFadden Act, 227
Medium of exchange, 49–50
Merger review policy, 239–240
Micro hedge, 306
Microprudential supervision, 263–264
Microsoft, 226
Mining, 54
Mitsubishi UFJ Financial Group, 236*t*
Modern Monetary Theory, 513*b*
Monetarists, 620
Monetary aggregates
Bank of Canada, 56–59
components of, 56*t*
growth rates, 60, 61*f*
money as weighted aggregate, 59–60
money supply definitions, 57–59, 57*f*–58*f*
Monetary base, 335, 356, 361–362
money supply, 358–362
in United States, 376*f*

Monetary conditions, 391
Monetary liabilities, 356
Monetary policy, 10–12
  aggregate demand shock, 593–594, 594*f,* 595*f*
  Bank of Canada and, 335, 390–395, 542–543
  case for rules, 620–621
  causes of, 601–605
  constrained discretion, 622–623
  conventional monetary policy, 395–405
  credibility, 623–631
  discretion, 619–622
  economic growth, 427
  at effective lower bound, 605–614
  European Central Bank, 420–421
  Federal Reserve, 414–419
  foreign exchange market, 493–494
  foreign exchange markets, stability in, 428
  global financial crisis, 436–438
  high employment and output stability, 426–427
  implementation of, 381–387
  inflation, 600–601, 601*f*
  inflation targeting, 429–435
  interest-rate stability, 427
  international financial system, 493–494
  Japanese, 612–614, 613*f*
  Large Value Transfer System (LVTS), 382, 385*b*
  Lucas critique, 617–618
  nominal anchor, 425, 623–631
  nominal interest rates and, 393–395
  nonconventional monetary policy tools. *see* Nonconventional monetary policy tools
  overnight rate, 381
  policy instrument, 443–446, 443*f*
  policy rate, 383–384
  price stability, 424, 429
  stability of financial markets, 427
  stabilizing economic activity, 597–600
  stabilizing inflation, 597
  standing facilities, 386
  stock prices and, 148
  supply shock, 594–597, 596*f*–598*f*
  Taylor rule, 446–449
  time-inconsistency problem, 425–426, 619–620
  transmission mechanisms of. *see* Transmission mechanisms of monetary policy
  types of rules, 620
Monetary policy (MP) curve, 542, 544*f*
  definition of, 543
  federal funds rate, 546–547, 546*f*
  shifts in, 544–545, 545*f,* 547–548, 551, 554
  Taylor principle, 543
Monetary Policy Report, 333, 339
Monetary theory, 8, 505
  budget deficits and inflation, 511–514
  quantity theory of money, 505–511
Monetary unions, 489–490
Monetizing the debt, 512
Money, 14, 15
  aggregate price level and, 8*f*
  business cycles and, 7–8
  cheques, 52–53
  commodity money, 52
  definition of, 7, 48
  electronic payment, 53
  e-money, 53–55, 54*b*
  fiat money, 52
  fiscal policy, 11–12
  functions of, 49–51
  growth rate, 8*f,* 10*f,* 21
  inflation and, 8–10, 10*f*
  and interest rates, 10, 11*f,* 109–113
  measurement of, 55–60
  medium of exchange, 49–50
  monetary policy, 10–12
  payments system, evolution of, 52–55
  shifts in demand for, 106, 107*t*
  shifts in supply of, 106–107, 107*t*
  store of value, 51
  supply and demand in market for, 103–106
  unit of account, 50–51
Money centre banks, 284
Money market, 103*n*
  certificate of deposit (CD), 28
  commercial paper, 29
  definition of, 27
  Government of Canada Treasury Bills, 28
  instruments, 27–29, 27*f*
  overnight funds, 29
  rates, 27*b*
  repurchase agreements (repos), 29
Money market mutual funds, 41
  Bruce Bent and, 221*b*
  definition of, 220
Money multiplier, 371–377
  definition of, 371
  derive, 371–373
  intuition behind, 373–374
Money supply, 7, 48, 107
  Bank of Canada's balance sheet, 356–358
  banks, 355
  borrowed reserves, 369
  central bank, 355
  changes in, 108–109, 109*f*
  coronavirus crises, 375–377
  currency holdings changes, 370
  depositors, 355
  desired reserve ratio, 369–370
  excess reserves changes, 370
  foreign exchange interventions and, 477–479
  growth rate of, 111–113, 112*f,* 113*f*
  loans to financial institutions, 360–361
  monetary aggregates, 57–59, 57*f*–58*f*

Money supply (*Continued*)
monetary base, 358–362
money multiplier, 371–377
multiple deposit creation, 363–369
nonborrowed monetary base, 369
overview of, 370, 371*t*
shifts from deposits into currency, 359–360
Montreal Accord, 260
Montreal Exchange (MX), 26
Moral hazard, 170
collateral, 179–180
in debt contracts, 176–177
definition of, 170
in equity contracts, 176–177
financial intermediaries, 36–38
financial intermediation, 181
government safety net and, 192
net worth, 179–180
principal-agent problem, 176–177
restrictive covenants, 180–181
summary, 181
Morgan Stanley, 226*t*
Mortgage-backed securities, 30
Mortgage markets
agency problems in, 250–252
financial innovation in, 250
Mortgages, definition of, 30
Multilateral netting, 383
Multiple deposit creation, 367–368
banking system, 364–367, 366*t*
critique of simple model, 368–369
definition of, 363
single bank, 363–364
Municipal government bonds, 31
Municipals, 31
Mutual funds, 41, 169

N

National banks, 210*b*
National Central Banks (NCBs), 349
National Financial Services OmbudService (NFSO), 240
National Housing Act, 261
Natural rate of output, 427, 566
Natural rate of unemployment, 426, 566
Negative demand shock, 624–626
Negative supply shock, 582, 582*f*
Net Canadian incurrence of liabilities, 481
Net capital inflows, 482
Net exports, 481, 524
definition of, 562
exchange rate effects on, 638–640
Net income, 25
Net investment income, 481
Net stable funding ratio, 264
Net worth, 175
adverse selection, 175
moral hazard, 179–180
New York Stock Exchange, 32*b*
NINJA loans, 199
Nominal anchor, 425, 623–631
Nominal gross domestic product, 19, 632
Nominal interest rate, 87, 471
definition of, 81
real *vs.* nominal, 81–83, 83*f*
Nonaccelerating inflation rate of unemployment (NAIRU), 447
Nonactivists, 598, 600*b*
Nonbank loans, 165
Nonborrowed monetary base, 362, 369
Nonborrowed reserves (NBR), 388, 444*f*
Nonconventional monetary policy tools
vs. credit easing, 410–411
definition of, 406
forward guidance, 411–412
large-scale asset purchases, 408–409, 408*f*
liquidity provision, 406–408, 609
management of expectations, 611–612, 612*f*
negative interest rates on banks' deposits, 412–413
quantitative easing and, 410–411, 609–612
Nontradable, 459
Note issuance facilities, 295
Notice deposits, 274
Notional principal, 320

O

Off-balance-sheet activities, 194, 294–297
generation of fee income, 295
loan sales, 295
risk management techniques, 295–297
trading activities, 295–297
Oil Drillers' Bank (ODB), 323
Oil price shocks, 627
One-period valuation model, 144–145
Ontario Municipal Improvement Corporation, 31
Ontario Securities Commission (OSC), 198
Open interest, 307
Open market buyback operations, 335
Open market operations
Bank of Canada, 358–359
definition of, 358
Open market purchase, 358–359
Open market sale, 358, 359
Operating band
definition of, 384
overnight interest rate, 384–386, 384*f*, 387
Operating instrument, 443
Operational independence, 339
Opportunity cost, 104
Optimal forecast, 149
Options, 311–320
American options, 311
definition of, 311
European options, 311
factors affecting option premiums, 318–319
financial futures options, 311
futures options, 311, 314–317
stock options, 311–314
summary, 319–320
vs. futures contracts, 316*f*

Orderly Liquidation Authority, 265, 267–268
Organisation for Economic Co-operation and Development (OECD), 194–195
Originate-todistribute, 218
Output gap, 558, 559*f,* 568, 571–572, 572*f,* 593
Overdraft loans, 275
Overnight cash rate, 420
Overnight funds, 29
Overnight interest rate, 29, 445*f*
  Bank of Canada's operating procedures, 389–390
  definition of, 383
  demand and supply in market for reserves, 387, 387*f*
  demand curve, 387–388
  market equilibrium, 389
  operating band for, 384–386, 384*f,* 387
  supply curve, 388–389
  Taylor rule, 448*f*
Overnight rate, 381
Over-the-counter (OTC) market, 26

P

Paper currency, 52
Paper loss, 79
Partisan business cycles, 342*b*
Par value, 68
Payments system, evolution of, 52–55
Payment technology, 515
Payoff method, 190
Pension funds, 41
Perfect substitutes, 128
Perpetuity
  definition of, 74
  yield to maturity on, 74–75
Personal consumption expenditures (PCE), 20
  deflator, 20
Phillips curve theory, 447
Plain vanilla swap, 320
Planned expenditure, 523
Planned investment spending, 524
  business expectations and, 526–527
  fixed investment, 525
  inventory investment, 525–526
  investment function, 527
  and real interest rates, 526
PNC Financial Services Group, 226*t*
Policy instrument
  controllability, 445–446
  definition of, 443
  measurability, 445
  monetary policy, 443–446, 443*f*
  observability, 445
  predictable effect on goals, 446
Policy rate, 383
Policy Targets Agreement, 430
Policy trilemma, 488–489, 488*f*
Political business cycle, 341, 621*b*
Portfolio, 36
Portfolio theories of money demand, 516–518
  and Keynesian liquidity preference, 516–517
  risk, 517
  summary, 517, 518*t*
  wealth, 517
Portfolio theory, 87
Positive demand shock, 579*f,* 624
Positive supply shock, 582
Potential GDP, 558
Potential output, 427, 558, 566
Precautionary motive, 515
Preferred habitat theories, 131–135
Premium, 311
Present discounted value, 66, 67
Present value
  definition of, 66
  present discounted value, 67
  simple loan, 66
Price level, 106, 110, 561
  quantity theory of money and, 507–508
Price-level targeting, *vs.* inflation-rate targeting, 390
Primary credit, 414
Primary dealers, 358, 395
Primary markets, 25–26
Principal–agent problem, 176–177
  debt contracts, 178
  financial intermediation, 178
  government regulation to increase information, 177–178
  monitoring, 177
  off-balance-sheet activities, 296*b,* 297
Printing money, 513
Private-equity firm, 178
Producer price index (PPI), 561
Property and Casualty (P&C) insurance companies, 41
Provincial Bond Purchase Program, 407*b,* 408
Provincial bonds, 31
Provincial Money Market Purchase Program, 407*b,* 408
Provincial Notes Act, 211
Provincials, 31
Prudential supervision, 195–197
Public interest view, 343
Purchase-and-assumption method, 190
Pure arbitrage, 153
Put option, 313–314, 314*f*

Q

Quantitative easing
  asset purchases and, 609–611
  in channel/corridor system, 409*b*
  definition of, 408
  nonconventional monetary policy and, 609–612
  vs. credit easing, 410–411
Quantity theory of money, 505–511
  definition of, 507
  demand for money, 507
  determinants of velocity, 506–507
  equation of exchange, 506, 507
  inflation and, 508–509, 510*f,* 511*f*
  price level and, 507–508
  testing, 509
  velocity of money, 506
Quotas, 459

R

Random walk, 154–155, 155*b*
Rasminsky, Louis, 338
Rate of capital gain, 77
Rate of return, definition of, 76–77
Rational expectations, 149
Rational expectations revolution, 617
Real exchange rate, 457
Real gross domestic product, 20
Real interest rate, 393, 471, 526, 637
  calculation of, 82
  definition of, 81
  Fisher equation, 81
  net exports and, 528
  nominal *vs.* real, 81–83, 83*f*
  planned investment spending and, 526
Real money balances, 515
Real terms, 81
Receiver general auctions, 398–399
Recessions, 8, 557
Recognition lag, 599
Redemption, 31
Redeposits, 398
Reference rate, 383
Registered bonds, 31
Regulatory arbitrage, 195
Repos, 395
Repurchase agreements (repos), 29, 254
Reserve currency, 484
Reserves, 275–276, 280–283, 357
  demand and supply in market for, 387, 387*f*
Residential housing prices, 252–254, 252*f,* 253*b*
Residual claimant, 25, 144
Resolution authority, 265, 267–268
Restrictive covenants, 168
  credit risk, 290
  definition of, 168
  moral hazard, 180–181
Return on assets (ROA), 286
Return on equity (ROE), 286
Returns
  definition of, 76
  interest rates and, 76–81, 79*t*
Revaluation, 485
Reverse repos, 396
Reverses, 396
Reverse transactions, 420
Revolving underwriting facilities, 295
Risk, 36
  asset demands and, 88–89, 89*t*
  bond demand and, 95*t,* 96
  definition of, 88
  demand for money, 517
Risk and liquidity premium, 123
Risk-averse, 88
Risk lover, 88
Risk management
  credit derivatives, 323–325
  financial derivatives, 301
  financial futures contract, 304–311
  financial futures markets, 307–309
  hedging, 301–302
  interest-rate forward contracts, 302–303
  options, 311–320
  swaps, 320–322
Risk preferrer, 88
Risk premium, 119, 123
Risk sharing, 36
Risk structure of interest rates, 118, 119*f*
  default risk, 119–123, 120*f*
  income tax considerations, 124
  liquidity, 123
  summary, 124–125

S

Sale and Repurchase Agreements (SRAs), 396, 397
  advantages of, 397
  definition of, 396
Sarbanes-Oxley Act, 199
Sarbanes-Oxley Act of 2002, 32*b*
Schedule I banks, 224–225
Schedule II banks, 224–225, 236
Schedule III banks, 224–225, 236
Seasonal credit, 415
Secondary credit, 414
Secondary loan participation, 295
Secondary markets, 25–26
Secondary reserves, 276
Secular stagnation, 102–103
Secured debt, 168
Security
  bank balance sheet, 276
  Bank of Canada's balance sheet, 357
  definition of, 3
  direct finance and, 23
  government agency securities, 31
  primary market, 25–26
  secondary market, 26
Segmented markets theory, 130–131
Seigniorage, 211
Seignorage, 500
Self-correcting mechanism, 578, 607–608, 608*f*
Senior tranche, 251*b*
Settlement balances, 275, 357
Settlement balances management, 398–400
Shadow banking system, 213–223
  adjustable-rate mortgages, 215
  avoidance of existing regulations, 219–221
  decline of traditional banking, 221–223
  definition of, 214
  financial derivatives, 215
  financial engineering, 214
  financial innovation, 221–223
  global financial crisis, 254–255, 255*f*
  information technology, 215–218
  interest-rate volatility, 214–215
  securitization and, 218–219
Short in a call, 312
Short position, 301, 304
Short sales, 160

Short term, definition of, 25
Simple deposit multiplier, 367
Simple interest rate, 66
Simple loan, 66
  yield to maturity, 69–70
Single bank, 363–364
Smart card, 53
Smart money, 154
Sovereign loans, 235
Special Purchase and Resale Agreements (SPRAs), 395–397
  advantages of, 397
  mechanics of, 396*f*
Special purpose vehicle (SPV), 251*b*
Speculative attack, 486
Speculative motive, 515
Spot exchange rate, 455
Spot transactions, 455
Stagflation, 582
Standardized gap analysis, 293
Standing lending facility, 402
  definition of, 402
  expansion, 406
  lender of last resort, 403–404, 404*f*
  operations, 402–405
Standing liquidity facilities, 386
Standing Term Liquidity Facility, 406, 407*b*
State banks, 210*b*
State-owned banks, 184
Sterilized foreign exchange intervention, 479, 480
Sticky, 569
Sticky prices, 637
Stockholders, definition of, 143
Stock market, 3–5, 5*f,* 155
  common stock, 3
  crash, 248
  lemons problem in, 171
  stock prices, 4, 5*f*
Stock options, 311–314
Stock prices, 248*f*
  efficient market hypothesis, 154–155, 157
  financial crisis, 249
  Gordon growth model, 147
  monetary policy and, 148
Stocks, 23, 49, 165
  capital market instruments, 29
  secondary market, 26
Stored-value card, 53
Store of value, 51
Stress tests, 198, 265
Strike price, 311
Structural unemployment, 426
Structured credit products, 250
Subprime mortgage market, 219
Supply curve, 91–92, 388–389
  for domestic assets, 463
Supply shock, 568, 594–597, 596*f*–598*f*
Supply-side economics, 427
Swaps
  currency swaps, 320
  definition of, 320
  with Exchange Fund Account, 399–400
  interest-rate swaps, 320–322
Swaptions, 320
Sweep accounts, 220–221
Swiss central bank, 267
Switzerland, 622*b*
Systemically important financial institutions (SIFIs), 265
Systemic risk, 382

T

T-accounts, 277–278, 358, 363
Target financing rate, 420
Tariffs, 459
Taxes
  aggregate demand, 527–528, 564
  IS curve, 534–535
Taylor, John, 446
Taylor principle, 447
TD Group US, 226*t*
Term deposit receipts, 28
Term notes, 28
Term structure of interest rates, 118, 618–619
  expectations theory, 127–130
  forecast interest rates, 137–139
  liquidity premium theory, 131–135
  preferred habitat theories, 131–135
  segmented markets theory, 130–131
Theory of bureaucratic behaviour, 343
Theory of efficient capital markets, 151
Theory of portfolio choice, 89
Theory of purchasing power parity (PPP), 457–461
  Big Macs and, 459–461, 460*t*
  evidence on, 458–459
  real exchange rate, 457
Theory of rational expectations, 143, 148
  adaptive expectations, 149
  efficient market hypothesis, 151
  formal statement of, 150
  implications of, 151
  optimal forecast, 149
  rationale of, 150–151
Three-tiered ownership regime, 238
Time-inconsistency problem, 425–426
Tobin's *q* theory, 640
Too-big-to-fail problem, 192–193
  Dodd-Frank, 267
  financial institutions, 266
  higher capital requirements, 266–267
Toronto Stock Exchange (TSX), 26
Toronto Stock Exchange (TSX) Venture Exchange, 26
Towers, Graham, 337
Toyota, 32*b*
Trade balance on goods and services, 481
Transaction costs, 34–35, 49, 168–169
  economies of scale, 169
  expertise, 169
  financial intermediaries reduce, 169
  in financial markets, 168–169
Transactions motive, 515
Transition countries, 182

Transmission mechanisms of monetary policy
credit channels, 644–645
credit view, 641–644
definition of, 637
exchange rate effects on net exports, 638–640
lessons for monetary policy, 645–648
Tobin's *q* theory, 640
traditional interest-rate channels, 637–638, 639*f*
wealth effects, 640–641
Treasury bills, 28
Troubled Asset Relief Program (TARP), 258
Trust companies, 39, 231
Trustees, 231
Tyranny of collateral, 183*b*

U

Unanticipated price level channel, 643
Uncertainty, 246
Underwrite, 42
Underwriting securities, 26
Unemployment rate, 7, 559, 561
Unexploited profit opportunities, 153
Unfavourable supply shock, 582
United States
AD/AS analysis, 585–587, 585*f,* 586*f*
banking industry, 225–227
capital markets, 32*b*
credit spreads in, 257*f*
current account deficit, 482*b*
fiscal stimulus, 535–536
excess reserves and currency ratio, 376*f*
Great Depression in, 248–250
market for reserves in, 415
monetary base in, 376*f*
negative supply shock, 583–584, 584*f*
Unit of account, 50–51
Universal banking, 230
Unsecured debt, 168
Unsterilized foreign exchange intervention, 478–480, 480*f*
U.S. Bankcorp, 226*t*

V

Value-at-risk (VaR), 198
Vault cash, 275, 357
Venture-capital firm, 178
Vietnam War Buildup, 533–531
Virtual bank, 217
Volcker rule, 265–266, 268

W

Wealth, 49
asset demands and, 88, 89*t*
bond demand and, 94, 95*t,* 96
definition of, 88
demand for money, 517
transmission mechanisms of monetary policy, 640–641
Wells Fargo & Co., 226*t*
Winnipeg Commodity Exchange, 26
World Bank, 483
World stock markets, 33
World Trade Organization (WTO), 483

Y

Yield curve, 125, 126*b*
definition of, 125
evidence on, 135
for forecast business cycles and inflation, 135*b*
for Government of Canada Bonds, 136*f*
interpretation, 136–137
liquidity premium theory, 134*f*
summary, 135–136
Yield to maturity, 65
coupon bond, 72–75, 73*t*
definition of, 69
discount bond, 75–76
fixed-payment loan, 70–71
simple loan, 69–70

Z

Zero-coupon bond, 69
Zimbabwean hyperinflation, 514